



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)	A2	(11) International Publication Number: WO 00/20601 (43) International Publication Date: 13 April 2000 (13.04.00)
---	-----------	--

(21) International Application Number: PCT/US99/22886

(22) International Filing Date: 1 October 1999 (01.10.99)

(30) Priority Data:

60/102,748	2 October 1998 (02.10.98)	US
60/123,810	11 March 1999 (11.03.99)	US
60/139,650	17 June 1999 (17.06.99)	US

(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive, Hayward, CA 94545 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US). CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).

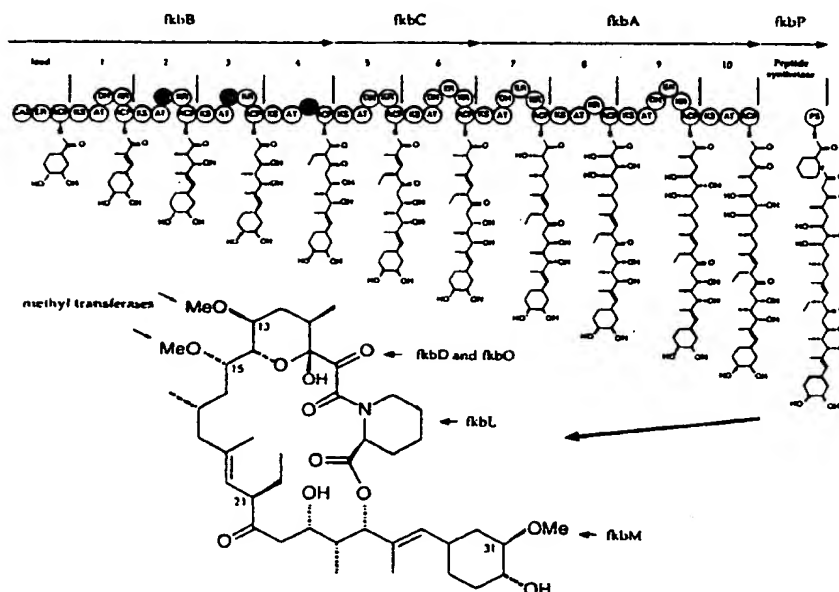
(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).

(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SI, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



(57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NI	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA
CONSTRUCTS THEREFOR

5

Field of the Invention

The present invention relates to polyketides and the polyketide synthase (PKS) enzymes that produce them. The invention also relates generally to genes encoding PKS enzymes and to recombinant host cells containing such genes and in which expression of such genes leads to the production of polyketides. The present invention also relates to
10 compounds useful as medicaments having immunosuppressive and/or neurotrophic activity. Thus, the invention relates to the fields of chemistry, molecular biology, and agricultural, medical, and veterinary technology.

Background of the Invention

15 Polyketides are a class of compounds synthesized from 2-carbon units through a series of condensations and subsequent modifications. Polyketides occur in many types of organisms, including fungi and mycelial bacteria, in particular, the actinomycetes. Polyketides are biologically active molecules with a wide variety of structures, and the class encompasses numerous compounds with diverse activities. Tetracycline,
20 erythromycin, epothilone, FK-506, FK-520, narbomycin, picromycin, rapamycin, spinocyn, and tylosin are examples of polyketides. Given the difficulty in producing polyketide compounds by traditional chemical methodology, and the typically low production of polyketides in wild-type cells, there has been considerable interest in finding improved or alternate means to produce polyketide compounds.

25 This interest has resulted in the cloning, analysis, and manipulation by recombinant DNA technology of genes that encode PKS enzymes. The resulting technology allows one to manipulate a known PKS gene cluster either to produce the polyketide synthesized by that PKS at higher levels than occur in nature or in hosts that otherwise do not produce the polyketide. The technology also allows one to produce
30 molecules that are structurally related to, but distinct from, the polyketides produced from known PKS gene clusters. See, e.g., PCT publication Nos. WO 93/13663; 95/08548; 96/40968; 97/02358; 98/27203; and 98/49315; United States Patent Nos. 4,874,748; 5,063,155; 5,098,837; 5,149,639; 5,672,491; 5,712,146; 5,830,750; and 5,843,718; and Fu *et al.*, 1994, *Biochemistry* 33: 9321-9326; McDaniel *et al.*, 1993, *Science* 262: 1546-1550; and Rohr, 1995, *Angew. Chem. Int. Ed. Engl.* 34(8): 881-888,
35 each of which is incorporated herein by reference.

Polyketides are synthesized in nature by PKS enzymes. These enzymes, which are complexes of multiple large proteins, are similar to the synthases that catalyze condensation of 2-carbon units in the biosynthesis of fatty acids. PKSs catalyze the biosynthesis of polyketides through repeated, decarboxylative Claisen condensations between acylthioester building blocks. The building blocks used to form complex polyketides are typically acylthioesters, such as acetyl, butyryl, propionyl, malonyl, hydroxymalonyl, methylmalonyl, and ethylmalonyl CoA. Other building blocks include amino acid like acylthioesters. PKS enzymes that incorporate such building blocks include an activity that functions as an amino acid ligase (an AMP ligase) or as a non-ribosomal peptide synthetase (NRPS). Two major types of PKS enzymes are known; these differ in their composition and mode of synthesis of the polyketide synthesized. These two major types of PKS enzymes are commonly referred to as Type I or "modular" and Type II "iterative" PKS enzymes.

In the Type I or modular PKS enzyme group, a set of separate catalytic active sites (each active site is termed a "domain", and a set thereof is termed a "module") exists for each cycle of carbon chain elongation and modification in the polyketide synthesis pathway. The typical modular PKS is composed of several large polypeptides, which can be segregated from amino to carboxy termini into a loading module, multiple extender modules, and a releasing (or thioesterase) domain. The PKS enzyme known as 6-deoxyerythronolide B synthase (DEBS) is a Type I PKS. In DEBS, there is a loading module, six extender modules, and a thioesterase (TE) domain. The loading module, six extender modules, and TE of DEBS are present on three separate proteins (designated DEBS-1, DEBS-2, and DEBS-3, with two extender modules per protein). Each of the DEBS polypeptides is encoded by a separate open reading frame (ORF) or gene; these genes are known as *eryAI*, *eryAII*, and *eryAIII*. See Caffrey *et al.*, 1992, *FEBS Letters* 304: 205, and U.S. Patent No. 5,824,513, each of which is incorporated herein by reference.

Generally, the loading module is responsible for binding the first building block used to synthesize the polyketide and transferring it to the first extender module. The loading module of DEBS consists of an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Another type of loading module utilizes an inactivated ketosynthase (KS) domain and AT and ACP domains. This inactivated KS is in some instances called KS^Q, where the superscript letter is the abbreviation for the amino acid, glutamine, that is present instead of the active site cysteine required for ketosynthase activity. In other PKS enzymes, including the FK-506 PKS, the loading module

incorporates an unusual starter unit and is composed of a CoA ligase like activity domain. In any event, the loading module recognizes a particular acyl-CoA (usually acetyl or propionyl but sometimes butyryl or other acyl-CoA) and transfers it as a thiol ester to the ACP of the loading module.

5 The AT on each of the extender modules recognizes a particular extender-CoA (malonyl or alpha-substituted malonyl, i.e., methylmalonyl, ethylmalonyl, and 2-hydroxymalonyl) and transfers it to the ACP of that extender module to form a thioester. Each extender module is responsible for accepting a compound from a prior module, binding a building block, attaching the building block to the compound from the prior
10 module, optionally performing one or more additional functions, and transferring the resulting compound to the next module.

Each extender module of a modular PKS contains a KS, AT, ACP, and zero, one, two, or three domains that modify the beta-carbon of the growing polyketide chain. A typical (non-loading) minimal Type I PKS extender module is exemplified by extender
15 module three of DEBS, which contains a KS domain, an AT domain, and an ACP domain. These three domains are sufficient to activate a 2-carbon extender unit and attach it to the growing polyketide molecule. The next extender module, in turn, is responsible for attaching the next building block and transferring the growing compound to the next extender module until synthesis is complete.

20 Once the PKS is primed with acyl- and malonyl-ACPs, the acyl group of the loading module is transferred to form a thiol ester (trans-esterification) at the KS of the first extender module; at this stage, extender module one possesses an acyl-KS and a malonyl (or substituted malonyl) ACP. The acyl group derived from the loading module is then covalently attached to the alpha-carbon of the malonyl group to form a carbon-
25 carbon bond, driven by concomitant decarboxylation, and generating a new acyl-ACP that has a backbone two carbons longer than the loading building block (elongation or extension).

The polyketide chain, growing by two carbons each extender module, is sequentially passed as covalently bound thiol esters from extender module to extender
30 module, in an assembly line-like process. The carbon chain produced by this process alone would possess a ketone at every other carbon atom, producing a polyketone, from which the name polyketide arises. Most commonly, however, additional enzymatic activities modify the beta keto group of each two carbon unit just after it has been added to the growing polyketide chain but before it is transferred to the next module.

Thus, in addition to the minimal module containing KS, AT, and ACP domains necessary to form the carbon-carbon bond, and as noted above, other domains that modify the beta-carbonyl moiety can be present. Thus, modules may contain a ketoreductase (KR) domain that reduces the keto group to an alcohol. Modules may also contain a KR domain plus a dehydratase (DH) domain that dehydrates the alcohol to a double bond. Modules may also contain a KR domain, a DH domain, and an enoylreductase (ER) domain that converts the double bond product to a saturated single bond using the beta carbon as a methylene function. An extender module can also contain other enzymatic activities, such as, for example, a methylase or dimethylase activity.

After traversing the final extender module, the polyketide encounters a releasing domain that cleaves the polyketide from the PKS and typically cyclizes the polyketide. For example, final synthesis of 6-dEB is regulated by a TE domain located at the end of extender module six. In the synthesis of 6-dEB, the TE domain catalyzes cyclization of the macrolide ring by formation of an ester linkage. In FK-506, FK-520, rapamycin, and similar polyketides, the TE activity is replaced by a RapP (for rapamycin) or RapP like activity that makes a linkage incorporating a pipecolate acid residue. The enzymatic activity that catalyzes this incorporation for the rapamycin enzyme is known as RapP, encoded by the *rapP* gene. The polyketide can be modified further by tailoring enzymes; these enzymes add carbohydrate groups or methyl groups, or make other modifications, i.e., oxidation or reduction, on the polyketide core molecule. For example, 6-dEB is hydroxylated at C-6 and C-12 and glycosylated at C-3 and C-5 in the synthesis of erythromycin A.

In Type I PKS polypeptides, the order of catalytic domains is conserved. When all beta-keto processing domains are present in a module, the order of domains in that module from N-to-C-terminus is always KS, AT, DH, ER, KR, and ACP. Some or all of the beta-keto processing domains may be missing in particular modules, but the order of the domains present in a module remains the same. The order of domains within modules is believed to be important for proper folding of the PKS polypeptides into an active complex. Importantly, there is considerable flexibility in PKS enzymes, which allows for the genetic engineering of novel catalytic complexes. The engineering of these enzymes is achieved by modifying, adding, or deleting domains, or replacing them with those taken from other Type I PKS enzymes. It is also achieved by deleting, replacing, or adding entire modules with those taken from other sources. A genetically engineered

PKS complex should of course have the ability to catalyze the synthesis of the product predicted from the genetic alterations made.

Alignments of the many available amino acid sequences for Type I PKS enzymes has approximately defined the boundaries of the various catalytic domains. Sequence alignments also have revealed linker regions between the catalytic domains and at the N- and C-termini of individual polypeptides. The sequences of these linker regions are less well conserved than are those for the catalytic domains, which is in part how linker regions are identified. Linker regions can be important for proper association between domains and between the individual polypeptides that comprise the PKS complex. One can thus view the linkers and domains together as creating a scaffold on which the domains and modules are positioned in the correct orientation to be active. This organization and positioning, if retained, permits PKS domains of different or identical substrate specificities to be substituted (usually at the DNA level) between PKS enzymes by various available methodologies. In selecting the boundaries of, for example, an AT replacement, one can thus make the replacement so as to retain the linkers of the recipient PKS or to replace them with the linkers of the donor PKS AT domain, or, preferably, make both constructs to ensure that the correct linker regions between the KS and AT domains have been included in at least one of the engineered enzymes. Thus, there is considerable flexibility in the design of new PKS enzymes with the result that known polyketides can be produced more effectively, and novel polyketides useful as pharmaceuticals or for other purposes can be made.

By appropriate application of recombinant DNA technology, a wide variety of polyketides can be prepared in a variety of different host cells provided one has access to nucleic acid compounds that encode PKS proteins and polyketide modification enzymes. The present invention helps meet the need for such nucleic acid compounds by providing recombinant vectors that encode the FK-520 PKS enzyme and various FK-520 modification enzymes. Moreover, while the FK-506 and FK-520 polyketides have many useful activities, there remains a need for compounds with similar useful activities but with better pharmacokinetic profile and metabolism and fewer side-effects. The present invention helps meet the need for such compounds as well.

Summary of the Invention

In one embodiment, the present invention provides recombinant DNA vectors that encode all or part of the FK-520 PKS enzyme. Illustrative vectors of the invention include cosmid pKOS034-120, pKOS034-124, pKOS065-C31, pKOS065-C3,

pKOS065-M27, and pKOS065-M21. The invention also provides nucleic acid compounds that encode the various domains of the FK-520 PKS, i.e., the KS, AT, ACP, KR, DH, and ER domains. These compounds can be readily used, alone or in combination with nucleic acids encoding other FK-520 or non-FK-520 PKS domains, as intermediates in the construction of recombinant vectors that encode all or part of PKS enzymes that make novel polyketides.

The invention also provides isolated nucleic acids that encode all or part of one or more modules of the FK-520 PKS, each module comprising a ketosynthase activity, an acyl transferase activity, and an acyl carrier protein activity. The invention provides an isolated nucleic acid that encodes one or more open reading frames of FK-520 PKS genes, said open reading frames comprising coding sequences for a CoA ligase activity, an NRPS activity, or two or more extender modules. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides isolated nucleic acids that encode all or a part of a PKS that contains at least one module in which at least one of the domains in the module is a domain from a non-FK-520 PKS and at least one domain is from the FK-520 PKS. The non-FK-520 PKS domain or module originates from the rapamycin PKS, the FK-506 PKS, DEBS, or another PKS. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides a method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector that encodes at least one module of a PKS, said module comprising at least one FK-520 PKS domain, and culturing said host cell under conditions such that said PKS is produced and catalyzes synthesis of said polyketide. In one aspect, the method is practiced with a *Streptomyces* host cell. In another aspect, the polyketide produced is FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-506 or rapamycin.

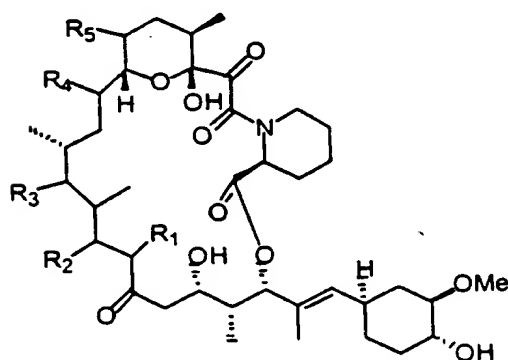
In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of ethylmalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require ethylmalonyl CoA for biosynthesis. The invention also provides recombinant nucleic acids that encode AT domains specific for ethylmalonyl CoA. Thus, the compounds of the invention can be

used to produce polyketides requiring ethylmalonyl CoA in host cells that otherwise are unable to produce such polyketides.

In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require 2-hydroxymalonyl CoA for biosynthesis. The invention also provides recombinant nucleic acids that encode AT domains specific for 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring 2-hydroxymalonyl CoA or 2-methoxymalonyl CoA in host cells that are otherwise unable to produce such polyketides.

In another embodiment, the invention provides a compound related in structure to FK-520 or FK-506 that is useful in the treatment of a medical condition. These compounds include compounds in which the C-13 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. Such compounds are less susceptible to the main *in vivo* pathway of degradation for FK-520 and FK-506 and related compounds and thus exhibit an improved pharmacokinetic profile. The compounds of the invention also include compounds in which the C-15 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. The compounds of the invention also include the above compounds further modified by chemical methodology to produce derivatives such as, but not limited to, the C-18 hydroxyl derivatives, which have potent neurotrophin but not immunosuppression activities.

Thus, the invention provides polyketides having the structure:



wherein, R₁ is hydrogen, methyl, ethyl, or allyl; R₂ is hydrogen or hydroxyl, provided that when R₂ is hydrogen, there is a double bond between C-20 and C-19; R₃ is hydrogen

or hydroxyl; R₄ is methoxyl, hydrogen, methyl, or ethyl; and R₅ is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506. The invention provides these compounds in purified form and in pharmaceutical compositions.

5 In another embodiment, the invention provides a method for treating a medical condition by administering a pharmaceutically efficacious dose of a compound of the invention. The compounds of the invention may be administered to achieve immunosuppression or to stimulate nerve growth and regeneration.

These and other embodiments and aspects of the invention will be more fully understood after consideration of the attached Drawings and their brief description below, together with the detailed description, examples, and claims that follow.

Brief Description of the Drawings

Figure 1 shows a diagram of the FK-520 biosynthetic gene cluster. The top line provides a scale in kilobase pairs (kb). The second line shows a restriction map with selected restriction enzyme recognition sequences indicated. K is *KpnI*; X is *XhoI*, S is *SacI*; P is *PstI*; and E is *EcoRI*. The third line indicates the position of FK-520 PKS and related genes. Genes are abbreviated with a one letter designation, i.e., C is *fkbc*. Immediately under the third line are numbered segments showing where the loading module (L) and ten different extender modules (numbered 1 - 10) are encoded on the various genes shown. At the bottom of the Figure, the DNA inserts of various cosmids of the invention (i.e., 34-124 is cosmid pKOS034-124) are shown in alignment with the FK-520 biosynthetic gene cluster.

Figure 2 shows the loading module (load), the ten extender modules, and the peptide synthetase domain of the FK-520 PKS, together with, on the top line, the genes that encode the various domains and modules. Also shown are the various intermediates in FK-520 biosynthesis, as well as the structure of FK-520, with carbons 13, 15, 21, and 31 numbered. The various domains of each module and subdomains of the loading module are also shown. The darkened circles showing the DH domains in modules 2, 3, and 4 indicate that the dehydratase domain is not functional as a dehydratase; this domain may affect the stereochemistry at the corresponding position in the polyketide. The substituents on the FK-520 structure that result from the action of non-PKS enzymes are also indicated by arrows, together with the types of enzymes or the genes that code for the enzymes that mediate the action. Although the methyltransferase is shown acting at the C-13 and C-15 hydroxyl groups after release of the polyketide from the PKS, the

methyltransferase may act on the 2-hydroxymalonyl substrate prior to or contemporaneously with its incorporation during polyketide synthesis.

Figure 3 shows a close-up view of the left end of the FK-520 gene cluster, which contains at least ten additional genes. The ethyl side chain on carbon 21 of FK-520 (Figure 2) is derived from an ethylmalonyl CoA extender unit that is incorporated by an ethylmalonyl specific AT domain in extender module 4 of the PKS. At least four of the genes in this region code for enzymes involved in ethylmalonyl biosynthesis. The polyhydroxybutyrate depolymerase is involved in maintaining hydroxybutyryl-CoA pools during FK-520 production. Polyhydroxybutyrate accumulates during vegetative growth and disappears during stationary phase in other *Streptomyces* (Ranade and Vining, 1993, *Can. J. Microbiol.* 39:377). Open reading frames with unknown function are indicated with a question mark.

Figure 4 shows a biosynthetic pathway for the biosynthesis of ethylmalonyl CoA from acetoacetyl CoA consistent with the function assigned to four of the genes in the FK-520 gene cluster shown in Figure 3.

Figure 5 shows a close-up view of the right-end of the FK-520 PKS gene cluster (and of the sequences on cosmid pKOS065-C31). The genes shown include *fk bD*, *fk bM* (a methyl transferase that methylates the hydroxyl group on C-31 of FK-520), *fk bN* (a homolog of a gene described as a regulator of cholesterol oxidase and that is believed to be a transcriptional activator), *fk bQ* (a type II thioesterase, which can increase polyketide production levels), and *fk bS* (a crotonyl-CoA reductase involved in the biosynthesis of ethylmalonyl CoA).

Figure 6 shows the proposed degradative pathway for tacrolimus (FK-506) metabolism.

Figure 7 shows a schematic process for the construction of recombinant PKS genes of the invention that encode PKS enzymes that produce 13-desmethoxy FK-506 and FK-520 polyketides of the invention, as described in Example 4, below.

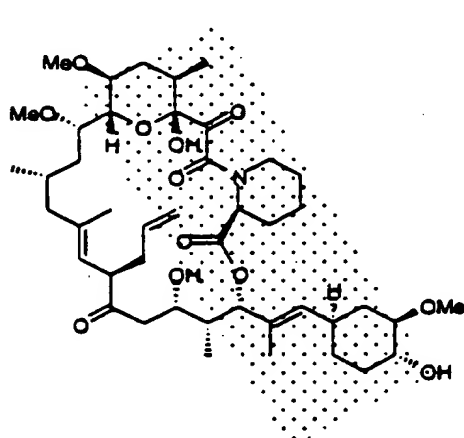
Figure 8, in Parts A and B, shows certain compounds of the invention preferred for dermal application in Part A and a synthetic route for making those compounds in Part B.

Detailed Description of the Invention

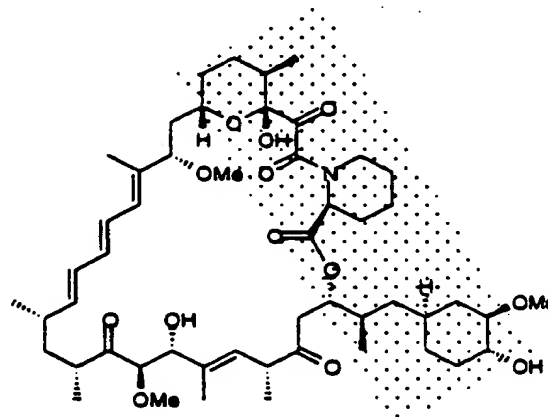
Given the valuable pharmaceutical properties of polyketides, there is a need for methods and reagents for producing large quantities of polyketides, as well as for producing related compounds not found in nature. The present invention provides such

methods and reagents, with particular application to methods and reagents for producing the polyketides known as FK-520, also known as ascomycin or L-683,590 (see Holt *et al.*, 1993, *JACS* 115:9925), and FK-506, also known as tacrolimus. Tacrolimus is a macrolide immunosuppressant used to prevent or treat rejection of transplanted heart,
5 kidney, liver, lung, pancreas, and small bowel allografts. The drug is also useful for the prevention and treatment of graft-versus-host disease in patients receiving bone marrow transplants, and for the treatment of severe, refractory uveitis. There have been additional reports of the unapproved use of tacrolimus for other conditions, including alopecia
10 universalis, autoimmune chronic active hepatitis, inflammatory bowel disease, multiple sclerosis, primary biliary cirrhosis, and scleroderma. The invention provides methods and reagents for making novel polyketides related in structure to FK-520 and FK-506, and structurally related polyketides such as rapamycin.

The FK-506 and rapamycin polyketides are potent immunosuppressants, with chemical structures shown below.



FK-506

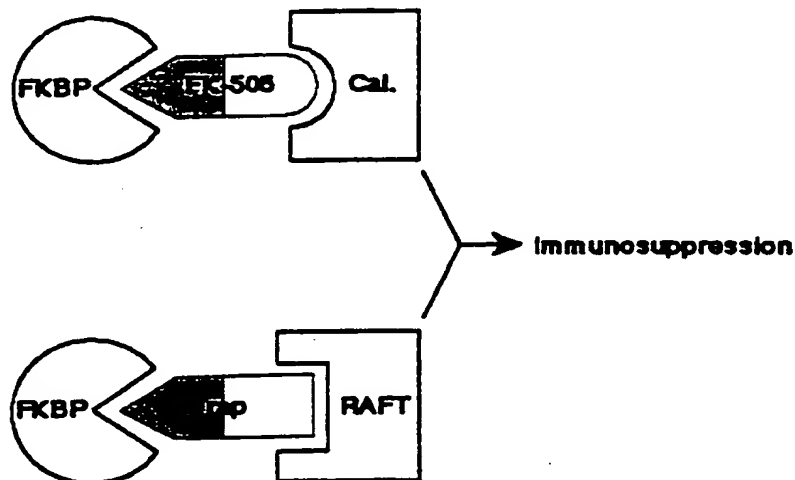


Rapamycin

15 FK-520 differs from FK-506 in that it lacks the allyl group at C-21 of FK-506, having instead an ethyl group at that position, and has similar activity to FK-506, albeit reduced immunosuppressive activity.

20 These compounds act through initial formation of an intermediate complex with protein "immunophilins" known as FKBP (FK-506 binding proteins), including FKBP-12. Immunophilins are a class of cytosolic proteins that form complexes with molecules such as FK-506, FK-520, and rapamycin that in turn serve as ligands for other cellular targets involved in signal transduction. Binding of FK-506, FK-520, and rapamycin to FKBP occurs through the structurally similar segments of the polyketide molecules,
25 known as the "FKBP-binding domain" (as generally but not precisely indicated by the

stippled regions in the structures above). The FK-506-FKBP complex then binds calcineurin, while the rapamycin-FKBP complex binds to a protein known as RAFT-1. Binding of the FKBP-polyketide complex to these second proteins occurs through the dissimilar regions of the drugs known as the "effector" domains.



The three component FKBP-polyketide-effector complex is required for signal transduction and subsequent immunosuppressive activity of FK-506, FK-520, and rapamycin. Modifications in the effector domains of FK-506, FK-520, and rapamycin that destroy binding to the effector proteins (calcineurin or RAFT) lead to loss of immunosuppressive activity, even though FKBP binding is unaffected. Further, such analogs antagonize the immunosuppressive effects of the parent polyketides, because they compete for FKBP. Such non-immunosuppressive analogs also show reduced toxicity (see Dumont *et al.*, 1992, *Journal of Experimental Medicine* 176, 751-760), indicating that much of the toxicity of these drugs is not linked to FKBP binding.

In addition to immunosuppressive activity, FK-520, FK-506, and rapamycin have neurotrophic activity. In the central nervous system and in peripheral nerves, immunophilins are referred to as "neuroimmunophilins". The neuroimmunophilin FKBP is markedly enriched in the central nervous system and in peripheral nerves. Molecules that bind to the neuroimmunophilin FKBP, such as FK-506 and FK-520, have the remarkable effect of stimulating nerve growth. *In vitro*, they act as neurotrophins, i.e., they promote neurite outgrowth in NGF-treated PC12 cells and in sensory neuronal cultures, and in intact animals, they promote regrowth of damaged facial and sciatic nerves, and repair lesioned serotonin and dopamine neurons in the brain. See Gold *et al.*, Jun. 1999, *J. Pharm. Exp. Ther.* 289(3): 1202-1210; Lyons *et al.*, 1994, *Proc. National Academy of Science* 91: 3191-3195; Gold *et al.*, 1995, *Journal of Neuroscience* 15:

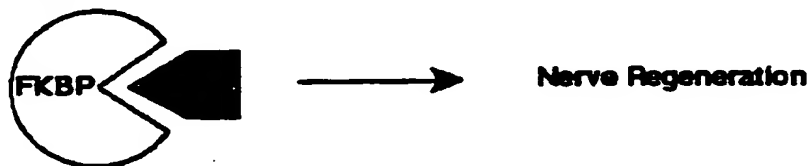
7509-7516; and Steiner *et al.*, 1997, *Proc. National Academy of Science* 94: 2019-2024.

Further, the restored central and peripheral neurons appear to be functional.

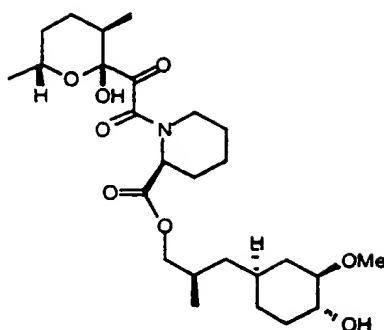
Compared to protein neurotrophic molecules (BNDF, NGF, etc.), the small-molecule neurotrophins such as FK-506, FK-520, and rapamycin have different, and often advantageous, properties. First, whereas protein neurotrophins are difficult to deliver to their intended site of action and may require intra-cranial injection, the small-molecule neurotrophins display excellent bioavailability; they are active when administered subcutaneously and orally. Second, whereas protein neurotrophins show quite specific effects, the small-molecule neurotrophins show rather broad effects.

Finally, whereas protein neurotrophins often show effects on normal sensory nerves, the small-molecule neurotrophins do not induce aberrant sprouting of normal neuronal processes and seem to affect damaged nerves specifically. Neuroimmunophilin ligands have potential therapeutic utility in a variety of disorders involving nerve degeneration (e.g. multiple sclerosis, Parkinson's disease, Alzheimer's disease, stroke, traumatic spinal cord and brain injury, peripheral neuropathies).

Recent studies have shown that the immunosuppressive and neurite outgrowth activity of FK-506, FK-520, and rapamycin can be separated; the neuroregenerative activity in the absence of immunosuppressive activity is retained by agents which bind to FKBP but not to the effector proteins calcineurin or RAFT. See Steiner *et al.*, 1997, *Nature Medicine* 3: 421-428.



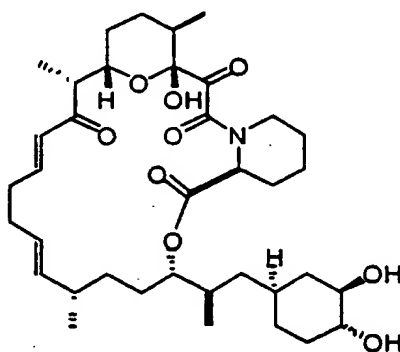
Available structure-activity data show that the important features for neurotrophic activity of rapamycin, FK-520, and FK-506 lie within the common, contiguous segments of the macrolide ring that bind to FKBP. This portion of the molecule is termed the "FKBP binding domain" (see VanDuyne *et al.*, 1993, *Journal of Molecular Biology* 229: 105-124.). Nevertheless, the effector domains of the parent macrolides contribute to conformational rigidity of the binding domain and thus indirectly contribute to FKBP binding.



"FKBP binding domain"

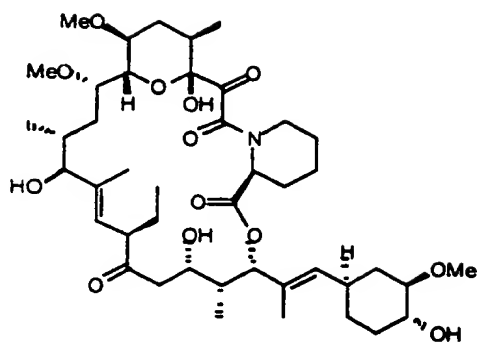
There are a number of other reported analogs of FK-506, FK-520, and rapamycin that bind to FKBP but not the effector protein calcineurin or RAFT. These analogs show effects on nerve regeneration without immunosuppressive effects.

- 5 Naturally occurring FK-520 and FK-506 analogs include the antascomycins, which are FK-506-like macrolides that lack the functional groups of FK-506 that bind to calcineurin (see Fehr *et al.*, 1996, *The Journal of Antibiotics* 49: 230-233). These molecules bind FKBP as effectively as does FK-506; they antagonize the effects of both FK-506 and rapamycin, yet lack immunosuppressive activity.

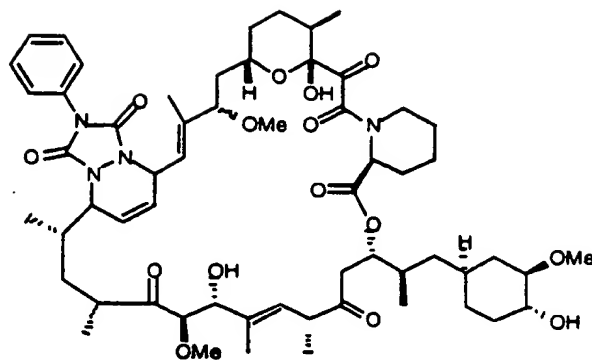


Antascomycin A

- 10 Other analogs can be produced by chemically modifying FK-506, FK-520, or rapamycin. One approach to obtaining neuroimmunophilin ligands is to destroy the effector binding region of FK-506, FK-520, or rapamycin by chemical modification. While the chemical modifications permitted on the parent compounds are quite limited,
- 15 some useful chemically modified analogs exist. The FK-520 analog L-685,818 (ED_{50} = 0.7 nM for FKBP binding; see Dumont *et al.*, 1992), and the rapamycin analog WAY-124,466 (IC_{50} = 12.5 nM; see Ocain *et al.*, 1993, *Biochemistry Biophysical Research Communications* 192: 1340-134693) are about as effective as FK-506, FK-520, and rapamycin at promoting neurite outgrowth in sensory neurons (see Steiner *et al.*, 1997).

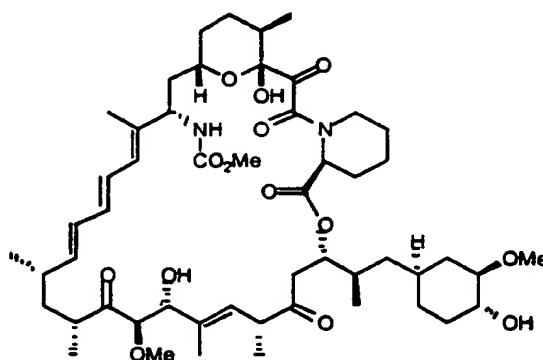


L-685,818



WAY-124,466

One of the few positions of rapamycin that is readily amenable to chemical modification is the allylic 16-methoxy group; this reactive group is readily exchanged by acid-catalyzed nucleophilic substitution. Replacement of the 16-methoxy group of rapamycin with a variety of bulky groups has produced analogs showing selective loss of immunosuppressive activity while retaining FKBP-binding (see Luengo *et al.*, 1995, *Chemistry & Biology* 2: 471-481). One of the best compounds, 1, below, shows complete loss of activity in the splenocyte proliferation assay with only a 10-fold reduction in binding to FKBP.



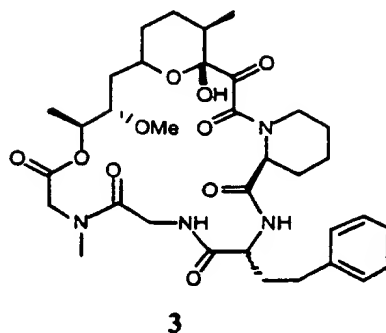
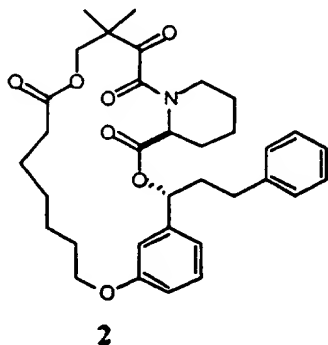
1

10

There are also synthetic analogs of FKBP binding domains. These compounds reflect an approach to obtaining neuroimmunophilin ligands based on "rationally designed" molecules that retain the FKBP-binding region in an appropriate conformation for binding to FKBP, but do not possess the effector binding regions. In one example, the ends of the FKBP binding domain were tethered by hydrocarbon chains (see Holt *et al.*, 1993, *Journal of the American Chemical Society* 115: 9925-9938); the best analog, 2, below, binds to FKBP about as well as FK-506. In a similar approach, the ends of the FKBP binding domain were tethered by a tripeptide to give analog 3, below, which binds

15

to FKBP about 20-fold poorer than FK-506. These compounds are anticipated to have neuroimmunophilin binding activity.



5 In a primate MPTP model of Parkinson's disease, administration of FKBP ligand GPI-1046 caused brain cells to regenerate and behavioral measures to improve. MPTP is a neurotoxin, which, when administered to animals, selectively damages nigral-striatal dopamine neurons in the brain, mimicking the damage caused by Parkinson's disease. Whereas, before treatment, animals were unable to use affected limbs, the FKBP ligand
10 restored the ability of animals to feed themselves and gave improvements in measures of locomotor activity, neurological outcome, and fine motor control. There were also corresponding increases in regrowth of damaged nerve terminals. These results demonstrate the utility of FKBP ligands for treatment of diseases of the CNS.

From the above description, two general approaches towards the design of non-
15 immunosuppressant, neuroimmunophilin ligands can be seen. The first involves the construction of constrained cyclic analogs of FK-506 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. The advantages of this approach are that the conformation of the analogs can be accurately modeled and predicted by computational methods, and the analogs closely resemble parent molecules that have
20 proven pharmacological properties. A disadvantage is that the difficult chemistry limits the numbers and types of compounds that can be prepared. The second approach involves the trial and error construction of acyclic analogs of the FKBP binding domain by conventional medicinal chemistry. The advantages to this approach are that the chemistry is suitable for production of the numerous compounds needed for such
25 interactive chemistry-bioassay approaches. The disadvantages are that the molecular types of compounds that have emerged have no known history of appropriate pharmacological properties, have rather labile ester functional groups, and are too conformationally mobile to allow accurate prediction of conformational properties.

The present invention provides useful methods and reagents related to the first
30 approach, but with significant advantages. The invention provides recombinant PKS

genes that produce a wide variety of polyketides that cannot otherwise be readily synthesized by chemical methodology alone. Moreover, the present invention provides polyketides that have either or both of the desired immunosuppressive and neurotrophic activities, some of which are produced only by fermentation and others of which are produced by fermentation and chemical modification. Thus, in one aspect, the invention provides compounds that optimally bind to FKBP but do not bind to the effector proteins. The methods and reagents of the invention can be used to prepare numerous constrained cyclic analogs of FK-520 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. Such compounds will show neuroimmunophilin binding (neurotrophic) but not immunosuppressive effects. The invention also allows direct manipulation of FK-520 and related chemical structures *via* genetic engineering of the enzymes involved in the biosynthesis of FK-520 (as well as related compounds, such as FK-506 and rapamycin); similar chemical modifications are simply not possible because of the complexity of the structures. The invention can also be used to introduce "chemical handles" into normally inert positions that permit subsequent chemical modifications.

Several general approaches to achieve the development of novel neuroimmunophilin ligands are facilitated by the methods and reagents of the present invention. One approach is to make "point mutations" of the functional groups of the parent FK-520 structure that bind to the effector molecules to eliminate their binding potential. These types of structural modifications are difficult to perform by chemical modification, but can be readily accomplished with the methods and reagents of the invention.

A second, more extensive approach facilitated by the present invention is to utilize molecular modeling to predict optimal structures *ab initio* that bind to FKBP but not effector molecules. Using the available X-ray crystal structure of FK-520 (or FK-506) bound to FKBP, molecular modeling can be used to predict polyketides that should optimally bind to FKBP but not calcineurin. Various macrolide structures can be generated by linking the ends of the FKBP-binding domain with "all possible" polyketide chains of variable length and substitution patterns that can be prepared by genetic manipulation of the FK-520 or FK-506 PKS gene cluster in accordance with the methods of the invention. The ground state conformations of the virtual library can be determined, and compounds that possess binding domains most likely to bind well to FKBP can be prepared and tested.

Once a compound is identified in accordance with the above approaches, the invention can be used to generate a focused library of analogs around the lead candidate, to "fine tune" the compound for optimal properties. Finally, the genetic engineering methods of the invention can be directed towards producing "chemical handles" that enable medicinal chemists to modify positions of the molecule previously inert to chemical modification. This opens the path to previously prohibited chemical optimization of lead compounds by time-proven approaches.

Moreover, the present invention provides polyketide compounds and the recombinant genes for the PKS enzymes that produce the compounds that have significant advantages over FK-506 and FK-520 and their analogs. The metabolism and pharmacokinetics of tacrolimus has been extensively studied, and FK-520 is believed to be similar in these respects. Absorption of tacrolimus is rapid, variable, and incomplete from the gastrointestinal tract (Harrison's Principles of Internal Medicine, 14th edition, 1998, McGraw Hill, 14, 20, 21, 64-67). The mean bioavailability of the oral dosage form is 27%, (range 5 to 65%). The volume of distribution (V₀D) based on plasma is 5 to 65 L per kg of body weight (L/kg), and is much higher than the V₀D based on whole blood concentrations, the difference reflecting the binding of tacrolimus to red blood cells. Whole blood concentrations may be 12 to 67 times the plasma concentrations. Protein binding is high (75 to 99%), primarily to albumin and alpha₁-acid glycoprotein. The half-life for distribution is 0.9 hour; elimination is biphasic and variable: terminal-11.3 hr (range, 3.5 to 40.5 hours). The time to peak concentration is 0.5 to 4 hours after oral administration.

Tacrolimus is metabolized primarily by cytochrome P450 3A enzymes in the liver and small intestine. The drug is extensively metabolized with less than 1% excreted unchanged in urine. Because hepatic dysfunction decreases clearance of tacrolimus, doses have to be reduced substantially in primary graft non-function, especially in children. In addition, drugs that induce the cytochrome P450 3A enzymes reduce tacrolimus levels, while drugs that inhibit these P450s increase tacrolimus levels. Tacrolimus bioavailability doubles with co-administration of ketoconazole, a drug that inhibits P450 3A. See, Vincent *et al.*, 1992, *In vitro* metabolism of FK-506 in rat, rabbit, and human liver microsomes: Identification of a major metabolite and of cytochrome P450 3A as the major enzymes responsible for its metabolism, *Arch. Biochem. Biophys.* 294: 454-460; Iwasaki *et al.*, 1993, Isolation, identification, and biological activities of oxidative metabolites of FK-506, a potent immunosuppressive macrolide lactone, *Drug Metabolism & Disposition* 21: 971-977; Shiraga *et al.*, 1994, Metabolism of FK-506, a

potent immunosuppressive agent, by cytochrome P450 3A enzymes in rat, dog, and human liver microsomes, *Biochem. Pharmacol.* 47: 727-735; and Iwasaki *et al.*, 1995, Further metabolism of FK-506 (Tacrolimus); Identification and biological activities of the metabolites oxidized at multiple sites of FK-506, *Drug Metabolism & Disposition* 23: 28-34. The cytochrome P450 3A subfamily of isozymes has been implicated as
5 important in this degradative process.

Structures of the eight isolated metabolites formed by liver microsomes are shown in Figure 6. Four metabolites of FK-506 involve demethylation of the oxygens on carbons 13, 15, and 31, and hydroxylation of carbon 12. The 13-demethylated (hydroxy)
10 compounds undergo cyclizations of the 13-hydroxy at C-10 to give MI, MVI and MVII, and the 12-hydroxy metabolite at C-10 to give I. Another four metabolites formed by oxidation of the four metabolites mentioned above were isolated by liver microsomes from dexamethasone treated rats. Three of these are metabolites doubly demethylated at the methoxy groups on carbons 15 and 31 (M-V), 13 and 31 (M-VI), and 13 and 15 (M-VII). The fourth, M-VIII, was the metabolite produced after demethylation of the 31-methoxy group, followed by formation of a fused ring system by further oxidation.
15 Among the eight metabolites, M-II has immunosuppressive activity comparable to that of FK-506, whereas the other metabolites exhibit weak or negligible activities. Importantly, the major metabolite of human, dog, and rat liver microsomes is the 13-demethylated and cyclized FK-506 (M-I).
20

Thus, the major metabolism of FK-506 proceeds via 13-demethylation followed by cyclization to the inactive M-I, this representing about 90% of the metabolic products after a 10 minute incubation with liver microsomes. Analogs of tacrolimus that do not possess a C-13 methoxy group would not be susceptible to the first and most important
25 biotransformation in the destructive metabolism of tacrolimus (i.e. cyclization of 13-hydroxy to C-10). Thus, a 13-desmethoxy analog of FK-506 should have a longer half-life in the body than does FK-506. The C-13 methoxy group is believed not to be required for binding to FKBP or calcineurin. The C-13 methoxy is not present on the identical position of rapamycin, which binds to FKBP with equipotent affinity as
30 tacrolimus. Also, analysis of the 3-dimensional structure of the FKBP-tacrolimus-calcineurin complex shows that the C-13 methoxy has no interaction with FKBP and only a minor interaction with calcineurin. The present invention provides C-13-desmethoxy analogs of FK-506 and FK-520, as well as the recombinant genes that encode the PKS enzymes that catalyze their synthesis and host cells that produce the
35 compounds.

These compounds exhibit, relative to their naturally occurring counterparts, prolonged immunosuppressive action *in vivo*, thereby allowing a lower dosage and/or reduced frequency of administration. Dosing is more predictable, because the variability in FK-506 dosage is largely due to variation of metabolism rate. FK-506 levels in blood
5 can vary widely depending on interactions with drugs that induce or inhibit cytochrome P450 3A (summarized in USP Drug Information for the Health Care Professional). Of particular importance are the numerous drugs that inhibit or compete for CYP 3A, because they increase FK-506 blood levels and lead to toxicity (Prograf package insert, Fujisawa □ US, Rev 4/97, Rec 6/97). Also important are the drugs that induce P450 3A
10 (e.g. Dexamethasone), because they decrease FK-506 blood levels and reduce efficacy. Because the major site of CYP 3A action on FK-506 is removed in the analogs provided by the present invention, those analogs are not as susceptible to drug interactions as the naturally occurring compounds.

Hyperglycemia, nephrotoxicity, and neurotoxicity are the most significant
15 adverse effects resulting from the use of FK-506 and are believed to be similar for FK-520. Because these effects appear to occur primarily by the same mechanism as the immunosuppressive action (i.e. FKBP-calcineurin interaction), the intrinsic toxicity of the desmethoxy analogs may be similar to FK-506. However, toxicity of FK-506 is dose related and correlates with high blood levels of the drug (Prograf package insert,
20 Fujisawa □ US, Rev 4/97, Rec 6/97). Because the levels of the compounds provided by the present invention should be more controllable, the incidence of toxicity should be significantly decreased with the 13-desmethoxy analogs. Some reports show that certain FK-506 metabolites are more toxic than FK-506 itself, and this provides an additional reason to expect that a CYP 3A resistant analog can have lower toxicity and a higher
25 therapeutic index.

Thus, the present invention provides novel compounds related in structure to FK-506 and FK-520 but with improved properties. The invention also provides methods for making these compounds by fermentation of recombinant host cells, as well as the recombinant host cells, the recombinant vectors in those host cells, and the recombinant
30 proteins encoded by those vectors. The present invention also provides other valuable materials useful in the construction of these recombinant vectors that have many other important applications as well. In particular, the present invention provides the FK-520 PKS genes, as well as certain genes involved in the biosynthesis of FK-520 in recombinant form.

FK-520 is produced at relatively low levels in the naturally occurring cells, *Streptomyces hygroscopicus* var. *ascomyceticus*, in which it was first identified. Thus, another benefit provided by the recombinant FK-520 PKS and related genes of the present invention is the ability to produce FK-520 in greater quantities in the recombinant host cells provided by the invention. The invention also provides methods for making novel FK-520 analogs, in addition to the desmethoxy analogs described above, and derivatives in recombinant host cells of any origin.

The biosynthesis of FK-520 involves the action of several enzymes. The FK-520 PKS enzyme, which is composed of the *fkfA*, *fkfB*, *fkfC*, and *fkfP* gene products, synthesizes the core structure of the molecule. There is also a hydroxylation at C-9 mediated by the P450 hydroxylase that is the *fkfD* gene product and that is oxidized by the *fkfO* gene product to result in the formation of a keto group at C-9. There is also a methylation at C-31 that is mediated by an O-methyltransferase that is the *fkfM* gene product. There are also methylations at the C-13 and C-15 positions by a methyltransferase believed to be encoded by the *fkfG* gene; this methyltransferase may act on the hydroxymalonyl CoA substrates prior to binding of the substrate to the AT domains of the PKS during polyketide synthesis. The present invention provides the genes encoding these enzymes in recombinant form. The invention also provides the genes encoding the enzymes involved in ethylmalonyl CoA and 2-hydroxymalonyl CoA biosynthesis in recombinant form. Moreover, the invention provides *Streptomyces hygroscopicus* var. *ascomyceticus* recombinant host cells lacking one or more of these genes that are useful in the production of useful compounds.

The cells are useful in production in a variety of ways. First, certain cells make a useful FK-520-related compound merely as a result of inactivation of one or more of the FK-520 biosynthesis genes. Thus, by inactivating the C-31 O-methyltransferase gene in *Streptomyces hygroscopicus* var. *ascomyceticus*, one creates a host cell that makes a desmethyl (at C-31) derivative of FK-520. Second, other cells of the invention are unable to make FK-520 or FK-520 related compounds due to an inactivation of one or more of the PKS genes. These cells are useful in the production of other polyketides produced by PKS enzymes that are encoded on recombinant expression vectors and introduced into the host cell.

Moreover, if only one PKS gene is inactivated, the ability to produce FK-520 or an FK-520 derivative compound is restored by introduction of a recombinant expression vector that contains the functional gene in a modified or unmodified form. The introduced gene produces a gene product that, together with the other endogenous and

functional gene products, produces the desired compound. This methodology enables one to produce FK-520 derivative compounds without requiring that all of the genes for the PKS enzyme be present on one or more expression vectors. Additional applications and benefits of such cells and methodology will be readily apparent to those of skill in the art after consideration of how the recombinant genes were isolated and employed in the construction of the compounds of the invention.

The FK-520 biosynthetic genes were isolated by the following procedure. Genomic DNA was isolated from *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) using the lysozyme/proteinase K protocol described in Genetic Manipulation of *Streptomyces* - A Laboratory Manual (Hopwood *et al.*, 1986). The average size of the DNA was estimated to be between 80 - 120 kb by electrophoresis on 0.3% agarose gels. A library was constructed in the SuperCos™ vector according to the manufacturer's instructions and with the reagents provided in the commercially available kit (Stratagene). Briefly, 100 µg of genomic DNA was partially digested with 4 units of *Sau*3A I for 20 min. in a reaction volume of 1 mL, and the fragments were dephosphorylated and ligated to SuperCos vector arms. The ligated DNA was packaged and used to infect log-stage XL1-BlueMR cells. A library of about 10,000 independent cosmid clones was obtained.

Based on recently published sequence from the FK-506 cluster (Motamedi and Shafiee, 1998, *Eur. J. Biochem.* 256: 528), a probe for the *fkfO* gene was isolated from ATCC 14891 using PCR with degenerate primers. With this probe, a cosmid designated pKOS034-124 was isolated from the library. With probes made from the ends of cosmid pKOS034-124, an additional cosmid designated pKOS034-120 was isolated. These cosmids (pKOS034-124 and pKOS034-120) were shown to contain DNA inserts that overlap with one another. Initial sequence data from these two cosmids generated sequences similar to sequences from the FK-506 and rapamycin clusters, indicating that the inserts were from the FK-520 PKS gene cluster. Two *Eco*RI fragments were subcloned from cosmids pKOS034-124 and pKOS034-120. These subclones were used to prepare shotgun libraries by partial digestion with *Sau*3AI, gel purification of fragments between 1.5 kb and 3 kb in size, and ligation into the pLitmus28 vector (New England Biolabs). These libraries were sequenced using dye terminators on a Beckmann CEQ2000 capillary electrophoresis sequencer, according to the manufacturer's protocols.

To obtain cosmids containing sequence on the left and right sides of the sequenced region described above, a new cosmid library of ATCC 14891 DNA was prepared essentially as described above. This new library was screened with a new *fkfM*

probe isolated using DNA from ATCC 14891. A probe representing the *fkpP* gene at the end of cosmid pKOS034-124 was also used. Several additional cosmids to the right of the previously sequenced region were identified. Cosmids pKOS065-C31 and pKOS065-C3 were identified and then mapped with restriction enzymes. Initial sequences from these cosmids were consistent with the expected organization of the cluster in this region. More extensive sequencing showed that both cosmids contained in addition to the desired sequences, other sequences not contiguous to the desired sequences on the host cell chromosomal DNA. Probing of additional cosmid libraries identified two additional cosmids, pKOS065-M27 and pKOS065-M21, that contained the desired sequences in a contiguous segment of chromosomal DNA. Cosmids pKOS034-124, pKOS034-120, pKOS065-M27, and pKOS065-M21 have been deposited with the American Type Culture Collection, Manassas, VA, USA. The complete nucleotide sequence of the coding sequences of the genes that encode the proteins of the FK-520 PKS are shown below but can also be determined from the cosmids of the invention deposited with the ATCC using standard methodology.

Referring to Figures 1 and 3, the FK-520 PKS gene cluster is composed of four open reading frames designated *fkpB*, *fkpC*, *fkpA*, and *fkpP*. The *fkpB* open reading frame encodes the loading module and the first four extender modules of the PKS. The *fkpC* open reading frame encodes extender modules five and six of the PKS. The *fkpA* open reading frame encodes extender modules seven, eight, nine, and ten of the PKS. The *fkpP* open reading frame encodes the NRPS of the PKS. Each of these genes can be isolated from the cosmids of the invention described above. The DNA sequences of these genes are provided below preceded by the following table identifying the start and stop codons of the open reading frames of each gene and the modules and domains contained therein.

	<u>Nucleotides</u>	<u>Gene or Domain</u>
	complement (412 - 1836)	<i>fkpW</i>
	complement (2020 - 3579)	<i>fkpV</i>
30	complement (3969 - 4496)	<i>fkpR2</i>
	complement (4595 - 5488)	<i>fkpR1</i>
	5601 - 6818	<i>fkpE</i>
	6808 - 8052	<i>fkpF</i>
	8156 - 8824	<i>fkpG</i>
35	complement (9122 - 9883)	<i>fkpH</i>
	complement (9894 - 10994)	<i>fkpI</i>
	complement (10987 - 11247)	<i>fkpJ</i>
	complement (11244 - 12092)	<i>fkpK</i>
	complement (12113 - 13150)	<i>fkpL</i>
40	complement (13212 - 23988)	<i>fkpC</i>

	complement (23992 - 46573)	<i>fk bB</i>
	46754 - 47788	<i>fk bO</i>
	47785 - 52272	<i>fk bP</i>
	52275 - 71465	<i>fk bA</i>
5	71462 - 72628	<i>fk bD</i>
	72625 - 73407	<i>fk bM</i>
	complement (73460 - 76202)	<i>fk bN</i>
	complement (76336 - 77080)	<i>fk bQ</i>
	complement (77076 - 77535)	<i>fk bS</i>
10	complement (44974 - 46573)	CoA ligase of loading domain
	complement (43777 - 44629)	ER of loading domain
	complement (43144 - 43660)	ACP of loading domain
	complement (41842 - 43093)	KS of extender module 1 (KS1)
	complement (40609 - 41842)	AT1
15	complement (39442 - 40609)	DH1
	complement (38677 - 39307)	KR1
	complement (38371 - 38581)	ACP1
	complement (37145 - 38296)	KS2
	complement (35749 - 37144)	AT2
20	complement (34606 - 35749)	DH2 (inactive)
	complement (33823 - 34480)	KR2
	complement (33505 - 33715)	ACP2
	complement (32185 - 33439)	KS3
	complement (31018 - 32185)	AT3
25	complement (29869 - 31018)	DH3 (inactive)
	complement (29092 - 29740)	KR3
	complement (28750 - 28960)	ACP3
	complement (27430 - 28684)	KS4
	complement (26146 - 27430)	AT4
30	complement (24997 - 26146)	DH4 (inactive)
	complement (24163 - 24373)	ACP4
	complement (22653 - 23892)	KS5
	complement (21420 - 22653)	AT5
	complement (20241 - 21420)	DH5
35	complement (19464 - 20097)	KR5
	complement (19116 - 19326)	ACP5
	complement (17820 - 19053)	KS6
	complement (16587 - 17820)	AT6
	complement (15438 - 16587)	DH6
40	complement (14517 - 15294)	ER6
	complement (13761 - 14394)	KR6
	complement (13452 - 13662)	ACP6
	52362 - 53576	KS7
	53577 - 54716	AT7
45	54717 - 55871	DH7
	56019 - 56819	ER7
	56943 - 57575	KR7
	57710 - 57920	ACP7
	57990 - 59243	KS8
50	59244 - 60398	AT8
	60399 - 61412	DH8 (inactive)
	61548 - 62180	KR8

62328 - 62537
 62598 - 63854
 63855 - 65084
 65085 - 66254
 5 66399 - 67175
 67299 - 67931
 68094 - 68303
 68397 - 69653
 69654 - 70985
 10 71064 - 71273

ACP8
 KS9
 AT9
 DH9
 ER9
 KR9
 ACP9
 KS10
 AT10
 ACP10

1 GATCTCAGGC ATGAAGTCCT CCAGGCGAGG CGCCGAGGTG GTGAACACCT CGCCGCTGCT
 61 TGTACGGACC ACTTCAGTCA GCGGCGATTG CGGAACCAAG TCATCCGGAA TAAAGGGCGG
 121 TTACAAGATC CTCACATTGC GCGACCGCCA GCATACGCTG AGTTGCCTCA GAGGCAAACC
 15 181 GAAAGGGCGC GGGCGGTCCG CACCAGGGCG GAGTACGCGA CGAGAGTGGC GCACCCGCGC
 241 ACCGTCACCT CTCTCCCCCG CCGGCGGGAT GCCCGGCGTG ACACGGTTGG GCTCTCCTCG
 301 ACGCTGAACA CCCGCGCGGT GTGGCGTCGG GGACACCGCC TGGCATCGGC CGGGTGACGG
 361 TACGGGGAGG GCGTACGGCG GCCGTGGCTC GTGCTCACGG CCGCCGGGCG GTCATCCGTC
 421 GAGACGGCAC TCGGCGAGCA GGGACGCCTG GTCGGCACCT GCGGGCCGGA CGACCGTGTG
 20 481 GTTCGCGGGC GGGCGGTGGC CGGTGGTGAG CCAGCTCTCC AGGGCGGTGA AGGCTGAGCG
 541 GTGACACGGC AGCAAAGGCC GGAGTCGGTC GGGGAAGGTG TCGACGAGGG CGTCGGTGTG
 601 CGTGCCGTCC TCGATGCGGT AGTAGCGGTA CCGGCCGCCA GGCCGCTGCC GGACATACGC
 661 GCGTACACGT CCGAGCCCGG GCGGCAGGCA GCAGCACGTC GAGAGTGCCT GGATGGTGAT
 721 CAGCGGCTTG CCGATACGAC CGGTCAACGC GATGCGTTCC ACGGCCGCGT GGACGCGGA
 25 781 GGAGCGGGTG GCGTAGTCGT AGTCGGCATC GCAGCCCGGG ACCGTCCCCG GGGCGCAATA
 841 CGGTGIGCCG GCTTCCTTCT CCCCATCGAA GCCGGGGTCG AACTCCTCGC GGTAGACGCG
 901 CTGCGTCAGA TCCCAGTAGA CCTCGTGGTG GTACGGCCAC AAGAACTCGG AGTCGGCCCG
 961 GAACCCGGCG CCGAGCAGCG CCTCGCGCGC CTGGCCGGCT GCGGGGCCGC CTGCCGCGTA
 1021 GGTGGGGTAG TCGCGCAGGG CCGCCGGCAG GAAGGTGAAG AGGTTGGGAC CCTCCGCGCG
 30 1081 CCACAGGGTG CCTTCCAGT CCACTCCTCC GTCGTACAGC TCGGGATGGT TCTCCAGCTG
 1141 CCAGCGCACG AGGTAGCCGC CGTTGGACAT CCCGGTGACC AGGGTGCGCT CGAGCGGCCG
 1201 GTGGTAGCGC TGGGCGACCG ACGCGCGGGC GGCCCGGGTC AGCTGGGTGA GGCGGGTGTT
 1261 CCACTCGGCG ACGGCGTCGC CCGGCCGGGA GCCATCACGG TAGAACGCGG GGCCTGTGTT
 1321 GCCCTTGTCG GTGGCGGCGT AGGCGTAACC GCGGGCGAGC ACCAGTCGG CGATGGCCCC
 35 1381 GTCGTTGGCG TACTGCTCGC GGTACCAGG GGTGCCGGCC ACGACAGGC CACCGTTCCA
 1441 GCGGTGCGGC AGCCGGATGA CGAACTGGGC GTCGTGGTTC CACCCGTGGT TGGTGTGGT
 1501 GGTGGAGGTG TCGGGGAAGT AGCCGTGAT CTGGATCCCG GGCACCTCCG TGGGAGTGGC
 1561 CAGGTTCTTG GCGGTCAGCC CTGCCCAGTC CGCCGGGTCG GTGTGGCCGG TGGCCGCCGT
 1621 TCCCGCCGTG GTCAGCTCGT CCAGGCAGTC GGCCTGCTGA CGTGCCGCCG CCGGGACACG
 40 1681 CAGCTGGGAC AGACGGGCGC AGTGACCGTC CCGGGCATCG GGAGCAGGCC GGGCCGTGGC
 1741 CCGTGAGGGG AGCAGGACGG CGACTGCGGC CAGGGTGAGA GCGCCGAGGC CGGTGCTCT
 1801 TCTCGGGGCC CGTCCGACAC CGAGGGGCAG AACCATGGAG AGCCTCCAGA CGTGCGGATG
 1861 GATGACGGAC TGGAGGCTAG GTGCGCACG GTGGAGACGA ACATGGGTGC GCCCGCATG
 1921 ACTGAGGCC CACAGAGGTG GGCCGCCGCC ATGACGGGCG CCGGACCGCG GCGCTCCCG
 45 1981 GGCGGIGCCC GCGGCCGCCA CCGGTTCCGG GTCCCGGGT CAGGGACAGG TGTCGTTCCG
 2041 GACCGTGAAG TAGCCGGTCG GCGACTCTTT CAAGGTGGTC GTGACGAAGG TGTTGTACAG
 2101 GCCCATGTTT TGGCCGGAGC CCTTGGCGTA GGTGTAACCG GCGTCTGTCG TGGCGCGGCC
 2161 CGCCTGGACG TGAGCGTAGT TGCCGGCGGT CCAGCAGACG GCCGTGGCAC CGGTGCTCTG
 2221 CGCGGTGACC GCGCCCGAGA GCGGTCCGGC CTTGCCGTCC GCGTCCCGGG CCGCGACCGC
 50 2281 GTAGGTGTGC GATGTGCCCG CCCTCAGGCC GGTGTCCGTG TACGACGTCG TGGCGGACGT
 2341 GGTGATCTGG GCACCGTCGC GGTGGACGGC GTAGTCGGTG GCGCCGTCGA CCGGTTTCCA
 2401 GGTGAGGCTG ATGGTGGTGT CCGTGGCGCC GGTGGCGGCC AGGCCGGACG GAGCGGGCAG
 2461 CGAACCGGGG TCGGAGGCGG ATCCGCTCAG GCCGAAGAAC TGCGTGATCC AGTAGCTGGA
 2521 ACAGATCGAG TCCAGGAAGT AGCGGCGGCC GGTGCTGCCG CACTGTGTCG CTCGGTGCC
 55 2581 GGGATCGACC GGGGTGCCGT GCCCGATGCC CCGCACCCGG TTACCTCCA CCGCCACCGA
 2641 TCCGTCCGCG GCCAGGTA CTCTGTGCCG GGTGGAGTTC GGGCCGATCA CCGAGGTACG
 2701 GTCCGGCGTC TGGGACACGC CGTGACACAG GGTCCACTGG TCGCGCAACT CGTCGGCGTT
 2761 GCGCGGCGCG ACGGTGGTGT CCTGTGCGCC GTGCCAGATG GCCACGCGCG GCCACGGGCC
 2821 CGACCACGAG GGGTAGCCGT CACGGACCCG CCGCGCCAC TGGTCCGCGG TCAGGTCCGT
 60 2881 CCGGGGGTTC ATGCACAGGT ACGCGCTGCT GACGTCCGTG GCACAGCCGA AGGGCAGGCC
 2941 GGCGACGACC GCGCCGGCCT GGAAGACGTC CGGATAGGTG GCGAGCATCA CCGACGTCAT

	3001	GGCACCGCCG	GCGGACAGCC	CGGTGATGTA	GGTGCGCTGG	GGGTCCGCGC	CGTAGGCGGA
	3061	GACGGTGTGA	GCGGCCATCT	GCCGGATCGA	CGCGGCTTCG	CCCTGGCCCC	TGCGGTTGTC
	3121	GCTGCTCTGG	AACCAGTTGA	AGCACCTGTT	CGCGTTGTTC	GACGACGTGG	TCTCGGCGAA
	3181	CACGAGCAGG	AAGCCATAGC	GGTCCGCGAA	TGAGAGCAGG	CCGGAGTTGT	CGGCGTAGCC
5	3241	CTGGGCGTCC	TGGGTGCAAC	CGTGCAGGGC	GAACACCACC	GCCGGCTCCG	CGGGCAGGGA
	3301	CGCGGGCCGG	TAGACGTACA	TGTTTCAGCCG	GCCCCGGGTT	GTGCCGAAGT	CCGCGACCTC
	3361	GGTCAGGTCC	GCCTTGGTCA	GACCGGGCTT	GGCCAGGCCC	GCCGCGGCGT	GGGCCGTCCG
	3421	CGCCGGGCGG	AGCAGGGCCG	CTCCGAGTAC	GAGGGCCACG	ACGGCCACGA	GACGGG"GAG
	3481	CACCCCCCGC	CGTCCCGGAC	GCGACAACGA	CCCGACCGGC	GGCGAGGAGG	AGAGGGGGAA
10	3541	CACGCGGGTG	AGGATTCCCC	GGAACGGCGG	CGGCTGCATG	GCGGCTCCCT	CAGTGTCTGT
	3601	GGGGGGACAC	GGAGGGCTCC	CTGACGTCGA	TCAGTGGGAG	CGCCCCGGTG	CCCGGCACCG
	3661	TAGGGGTGGT	TCAACCCGCA	ACGGTATGGC	CCGGAGCACC	ACACCCCGCA	CCGCGCGATG
	3721	TGCGCCCCGA	CGGATTGTGT	CGCCTTGCGG	AATCTGATAC	CCGGACGCGA	CGAACGCCCC
	3781	ACCCGACACG	GGTAGGGCGT	CATGGTGTCC	GACTCGGCCG	GTGGCCCTTG	CCTGCCCTGG
15	3841	ACGGACCGGG	CGTCGGCGGA	CCGGGCGTCG	GCGGGCTGGG	CGGTATGGCG	GCCGAGGACG
	3901	CCAGCCGCGT	GGGGCGGCGG	CGCCCAAGTG	CAGTACGCCG	ACCGTGGCCG	CGGGGAGGGC
	3961	CGGAACGGTC	AGTGCAGTCC	CGCGGCCCTG	CGGGACCGCT	CGTCCCAGAC	GGGTTCCACC
	4021	GCGGCGAACC	GGGGTCCGTG	TCCGCGGCGG	TAGACCATCA	GTGTCCGCTC	GAAGGTGATG
	4081	ACGATGACAC	CGTCCTGGTT	GTAGCCGATG	GTGCGCACGC	TGATGATGCC	TACGTACAGT
20	4141	CGGCTGGCGG	ACTCCCCGGT	GTTCAGGACC	TCCGACTGCG	AGTAGATGGT	GTCGCCCTCG
	4201	AAGACCGGGT	TCGGCAGCCT	GACCCGGTCC	CAGCCGAGGT	TGGCCATCAC	ATGCTGGGAG
	4261	ATGTTCGGTG	CGCTCTGCCC	GGTGACCAGG	GCGAGGGTGA	AGGTGGAGTC	CACCAAGGGC
	4321	TGTCCCCAGG	TGGTSCCCGC	CGAGTAGTGG	CGGTCTGAAGT	GCAGCGGCGC	GGTGTCTGTC
	4381	GTCAGGAGCG	TGAGCCAGGA	GTTGTGCGTC	TCCAGGACCG	TGCGGCCAG	GGGGTGGCGG
25	4441	TACACGTCGC	CGGTGGTGAA	GTCTCGAAG	TAGCGGCCCT	GCCAGCCCTC	GACCACAGCG
	4501	GTGCGGGTGG	CGTCCTGGTC	CGGGTTCTCA	GTCGTCATGG	CGCTCATTCT	GGGAAGTCCC
	4561	CGGTCCGCTG	TGAAATGCCG	AACCTTCACC	GGGCTCATAC	GTGCGGCGCA	TGAGCCCTGG
	4621	ACCGTACGTA	GTCTAGAAC	CTCGCCACCA	CTGGCGCGCG	TGGTCTCTCC	GCGAGTGTGA
30	4681	CCACGTCGAC	CGTGCGCCGC	GCCTGCGGGT	CGTCGAGCGG	CACGGCGACG	GCGTGGTCAC
	4741	CGGGCCCCGA	CGGGCTGCCG	GTGAGGGGGG	CGACGGCCAC	ACCGAGCCCG	GCGGCGACCA
	4801	GGGCCCCGAG	CGTGCTCAGC	TCGGTGCTCT	CCAGGACGAC	CCGCGGCACG	AATCCGGCCG
	4861	CGGCGCACAG	CCGGTCGGTG	ATCTGGCGCA	GTCCGAAGAC	CGGCTCCAGT	GCCACGAACG
	4921	CCTCATCGGC	CAGCTCCGCG	GTCCGCAACC	GGCGGCGTCT	GGCCAGCCGG	TGTCCGGGTG
	4981	GGACGAGCAG	GCACAGTGCC	TCGTCCCGCA	GTGGTGTCCA	CTCCACATCG	TCCCCGGCGG
35	5041	GTCTGTTGGT	GGTCAGCCCC	AGGTCCAGCC	TGCTGTTGCG	GACGTCGTCT	ACCACGGCGT
	5101	CGGCGGCGTC	GCCGCGCAGT	TGGAAGGTGG	TGCCGGGAGC	CAGCGGCGCG	TACCCGCGGA
	5161	GGAGGTGCGG	CACCAGCCAG	GTGCCGTAGG	AGTGCAGGAA	ACCCAGTGCC	ACGGTGCCGG
	5221	TGTCGGGGTC	GATCAGGGCG	GTGATGCGCT	GCTCGGCGCC	GGAGACCTCA	CTGATCGCGC
	5281	GCAGGGCGTG	GGCGCGGAAG	ACCTCGCCGT	ACTTGTGTAG	CCGGAGCCGG	TTCTGGTGCC
40	5341	GGTCGAACAG	CGGCACGCCC	ACTCGTCGCT	CCAGCCGCGG	GATGGCCCTG	GACAGGGTCG
	5401	GCTGGGAGAT	GTTGAGCCGT	TCCGCGGTGA	TCGTACAGTG	CTCGTGCTCG	GCCAAGGCCG
	5461	TGAACCACTG	CAACTCCCCT	ATCTCCATGC	AGGGACTATA	CGTACCGGGC	ATGGTCTTGG
	5521	CGAGGTTTCG	TCATTTACAC	GCGGCGGGGC	GCGGCGCCAC	AGTGAGTCTT	CACCAACCAG
	5581	GACCCCATGG	GAGGGACCCC	ATGTCCGAGC	CGCATCCTCG	CCCTGAACAG	GAACGCCCCG
45	5641	CCGGGCCCCCT	GTCCGGTCTG	CTCGTGGTTT	CTTTGGAGCA	GGCCGTGCGC	GCTCCGTTCTG
	5701	CCACCCGCCA	CCTGGCGGAC	CTGGGCGCCC	GTGTCATCAA	GATCGAACGC	CCCGGCAGCG
	5761	GCGACCTCGC	CCGCGGCTAC	GACCGCACGG	TGCGTGGCAT	GTCCAGCCAC	TTCGTCTGGC
	5821	TGAACCGGGG	GAAGGAGAGC	GTCCAGTCTG	ATGTGCGCTC	GCCGGAGGGC	AACCGGCACC
	5881	TGCACGCCTT	GGTGGACCGG	CCCGATGTCC	TGGTGCAGAA	TCTGGCACCC	GGCGCCGCGG
50	5941	GCCGCTTGGC	ATCGGCCACC	AGGTCTCTCG	GCGGAGCCAC	CGAGGCTGAT	CACCTGCGGA
	6001	CATATCCGGC	TACGGCAGTA	CCGGCTGCTA	CCGCGGACCG	CAAGGCGTAC	GACCTCCTGG
	6061	TCCAGTGCGA	AGCGGGGCTG	GTCTCCATCA	CCGGCACCCC	CGAGACCCCG	TCCAAGGTGG
	6121	GCCTGTCCAT	CGCGGACATC	TGTGCGGGGA	TGTACGCGTA	CTCCGGCATC	CTCACGGCCC
	6181	TGCTGAAGCG	GGCCCCGACC	GGCCGGGGCT	CGCAGTTGGA	GGTCTCGATG	CTCGAAGCCC
55	6241	TCGGTGAATG	GATGGGATAC	CCCGAGTACT	ACACGCGCTA	CGGCGGCGCC	GCTCCGGCCC
	6301	GCGCCGGCGC	CAGCCACGCG	ACGATCGCCC	CCTACGGCCC	GTTACACACG	CGCGACGGGC
	6361	AGACGATCAA	TCTCGGGCTC	CAGAACGAGC	GGGAGTGGGC	TTCCTTCTGC	GGTGTCTGTC
	6421	TACAACGCCC	CGGTCTCTGC	GACGACCCGC	GCTTTTCCGG	CAACGCCGAC	CGGGTGGCGC
	6481	ACCGCACCGA	GCTCGACGCC	CTGGTGAGCG	AGGTGACGGG	CACGCTCACC	GGCGAGGAAC
60	6541	TGGTGGCGCG	GCTGGAGGAG	GCGTCGATCG	CCTACGCACG	CCAGCGCACC	GTGCGGGAGT
	6601	TCAGCGAACA	CCCCCAACTG	CGTGACCGTG	GACGCTGGGC	TCCGTTTCAG	AGCCCGGTCTG
	6661	GTGCGCTGGA	GGGCCTGATC	CCCCCGGTCA	CCTTCCACGG	CGAGCACCCG	CGGCGGCTGG
	6721	GCCGGGTCCC	GGAGCTGGGC	GAGCATACCG	AGTCCGTCTT	GGCGTGCGCTG	GCCGCGCCCC
	6781	ACAGCGCCGA	CCGCGAAGAG	GCCGGCCATG	CCGAATGAAC	TCACCGGAGT	CCTGATCTCTG

	6841	GCCGCCGTGT	TCCTGCTCGC	CGGCGTACGG	GGGCTGAACA	TGGGCCTGCT	CGCGCTGGTC
	6901	GCCACCTTTC	TGCTCGGGGT	GGTTCGACTC	GACCGAACGC	CGGACGAGGT	GCTGGCGGGT
	6961	TTCCCCGCGA	GCATGTTCC	GGTGTGGTTC	GCCGTACAGT	TCCTCTTCGG	GATCGCCCGC
5	7021	GTCAACGGCA	CGGTGGACTG	GCTGGTACGT	GTGCGGGTGC	GGGCGGTGGG	GGCCCGGGTG
	7081	GGAGCCGTCC	CCTGGGTGCT	CTTCGGCCTG	GCGGCACTGC	TCTGCGCGAC	AGGCGCGGCC
	7141	TCGCCCGCGG	CGGTGGCGAT	CGTGGCGCCG	ATCAGCGTCG	CGTTCGCCGT	CAGGCACCGC
	7201	ATCGATCCGC	TGTACGCCGG	ACTGATGGCG	GTGAACGGGG	CCGCAGCCGG	CAGTTTCGCC
	7261	CCCTCCGGGA	TCCTGGGCGG	CATCGTCCAC	TCGGCGCTGG	AGAAGAACCA	TCTGCCCGTC
10	7321	AGCGGCGGGC	TGCTCTTCGC	AGGCACCTTC	GCCTTCAACC	TGGCGGTTCG	CGCGGTJTCA
	7381	TGGCTCGTCC	TCGGGCGCAG	GCGCCTCGAA	CCACATGACC	TGGACGAGGA	CACCGATCCC
	7441	ACGGAAGGGG	ACCCGGCTTC	CCGCCCCGGC	GCGGAACACG	TGATGACGCT	GACCGCGATG
	7501	GCCGCGCTGG	TGCTGGGAAC	CACGGTCCCT	TCCCTGGACA	CCGGCTTCCT	GGCCCTCACC
	7561	TTGGCCGCGT	TGCTGGCGCT	GCTCTTCCCG	CGCACCTCCC	AGCAGGCCAC	CAAGGAGATC
15	7621	GCCTGGCCCC	TGGTGTGCTG	GGTATGCGGG	ATCGTGACCT	ACGTCCGCCCT	ACGTCCGGAG
	7681	CTGGGCATCG	TGGACTCCCT	GGGGAAGATG	ATCGCGGCGA	TCGGCACCCC	GCTGCTGGCC
	7741	GCCCTCGTGA	TCTGCTACGT	GGGCGGTGTC	GTCTCGGCCT	TCGCTCGAC	CACCGGGATC
	7801	CTCGGTGCCC	TGATGCCGCT	GTCCGAGCCG	TTCTGAAGT	CCGGTGCCAT	CGGGACGACC
	7861	GCCATGGTGA	TGGCCCTGGC	GGCCGCGGGC	ACCGTGGTGG	ACGCGAGTCC	CTTCTCCACC
20	7921	AATGGTGCTC	TGGTGGTGGC	CAACGCTCCC	GAGCGGCTGC	GGCCCGGCGT	GTACCAGGGG
	7981	TTGCTGTGGT	GGGGCGCCGG	GGTGTGCGCA	TGTGCTCCCG	CGGCCGCGCT	GGCGCCCTTC
	8041	GTGGTGGCGT	GAGCGCAGCG	GAGCGGGAAT	CCCCTGGAGC	CCGTTTCCCG	TGCTGTGTCG
	8101	CTGACGTAGC	GTCAAGTCCA	CGTGCCGGGC	GGGCAGTACG	CCTAGCATGT	CGGGCATGGC
	8161	TAATCAGATA	ACCCTGTCCG	ACACGCTGCT	CGCTTACGTA	CGGAAGGTGT	CCCTGCGCGA
25	8221	TGACGAGGTG	CTGAGCCGGC	TGCGCGCGCA	GACGGCCGAG	CTGCCGGGCG	GTGGCGTACT
	8281	GCCGGTGCAG	GCCGAGGAGG	GACAGTTCCT	CGAGTTCCTG	GTGCGGTTGA	CCGGCGCGCG
	8341	TCAGGTGCTG	GAGATCGGGA	CGTACACCGG	CTACAGCAGC	CTCTGCCTGG	CCCGCGGATT
	8401	GGCGCCCGGG	GGCCGTGTGG	TGACGTGCGA	TGTCATGCCG	AAGTGGCCCG	AGGTGGGCGA
	8461	GCGGTACTGG	GAGGAGGCCG	GGGTTGCCGA	CCGGATCGAC	GTCCGGATCG	GCGACGJCCG
30	8521	GACCGTCCCT	ACCGGGCTGC	TCGACGAGGC	GGGCGCGGGG	CCGGAGTCGT	TCGACATGGT
	8581	GTTCAFCGAC	GCCGACAAGG	CCGGCTACCC	CGCCTACTAC	GAGGCGGCGC	TGCCGCTGGT
	8641	ACGCCGCGGC	GGGCTGATCG	TCGTGCAACA	CACGCTGTTC	TTCCGGCCGGG	TGGCCGACGA
	8701	AGCGGTGCAG	GACCCGGACA	CGGTTCGCGT	ACGCGAACTC	AACGCGGCAC	TGCGCGACGA
	8761	CGACCGGGTG	GACCTGGCGA	TGCTGACGAC	GGCCGACGGC	GTCACCCCTG	TCCGGAACCG
35	8821	GTGACCGGGG	CGATGTCCGG	GGCGGTTCAG	GTCAGCGTCG	TCCGGCGGGG	TCCGGCGGAG
	8881	GGCTCCAGAT	GCAGGCGTTC	GACGCCGGCG	GCGGAAGCGC	CCGCCACCTC	GGACACGCAG
	8941	GGGCGAGTCG	AGTCCGCGAA	GCCCGCGAAC	CGGTAGGCGA	TCTCCATCAT	GCGGTTGCGG
	9001	TCCGTACGCC	GGAAGTCCGC	CACCAGGTGC	GCCCCCGCGC	GGGCGCCCTG	GTCCGTGAGC
	9061	CAGTTCAGGA	TCGTGCGACC	GGCACCGAAC	GACACGACCC	GGCAGGACGT	GGCGAGCAGT
40	9121	TTCAGGTGCC	ACGTGACGGC	CTTCTTCTCC	AGCAGGATGA	TGCCGACGGG	GCCGTGCGGG
	9181	CCGAAGCGGT	CGCCCATGGT	GACGACGAGG	ACCTCATGGG	CGGGATCGGT	CGGACGCGC
	9241	GCAGGTGCGC	GTCGGAGTAG	TGCACGCCGG	TCGCGTTTCT	CTGGCTGGTC	CGCAGCGTCA
	9301	GTTCTCTGAC	GCGGCTGAGT	TCTCTCTCCC	CCGCGGGTGC	GATCGTTCAT	GAGAGGTGCA
	9361	GCGAGCGCAG	GAAGTCCTCG	TCGGGACCGG	AGTACGCCTC	CCGGGCTTGG	TCGCGCGCGA
45	9421	AACCCGCCTG	GTACATCAGG	CGGCGCCGAC	GCGAGTCGAC	CGTGGACACC	GGCGGGCTGA
	9481	ACTCCGGCAG	CGACAGGAGC	GTGGCCGCCT	GCTCGGCCGG	GTAGCACC	ACCTCGGGCA
	9541	GGTGGAAACG	CACCTCGGCA	CGCTCGGCGG	GCTGGTCTGC	GATGAACGCG	ATCGTGGTTC
	9601	GTGCGAAGTT	CAGCTCCGTG	GCGATCTCGC	GGACGGACTG	CGACTTCGGC	CCCATCCGA
	9661	TGCGGGCCAG	CACGAAGTAC	TCCGCCACAC	CGAGGCGTTC	CAGACGCTCC	CACGCGAGGT
50	9721	CGTGGTCGTT	CTTGCTCGCC	ACCGCCTGGA	GGATGCCGCG	GTCGTGAGC	GTGGTGATCA
	9781	CCTCGCGGAT	CTCGTCGGTG	AGGACCACCT	CGTCGTCTTC	CAGCACGGTG	CCCCGCCACA
	9841	AGGTGTTGTC	CAGGTCCCAG	ACCAGACACT	TGACAATGGT	CATGGCTGTC	CTCTCAAGCC
	9901	GGGAGCGCCA	GCGCGTGCTG	GGCCAGCATC	ACCCGGCACA	TCTCGTGCTC	GCCCTCGATG
	9961	ATCTCCATGA	GCTTGGCGTC	CGGTTACGCC	CGTTCGACGA	CGTGTCCCTC	TCTCGCGCCT
55	10021	GCCGACGCGA	GCACCTGTGC	GGCGGTTCGG	GCCCCGGCGG	CGGCTCGTTC	GGCGGCGACG
	10081	TGCTTGGCCA	GGATCGTCGC	GGGCACCATC	TCGGGCGAGC	CCTCGTCCCA	GTGGTTCGCTG
	10141	GCGTACTCGC	ACACGCGGGC	CGCGATCTGC	TCCGCGGTCC	ACAGGTTCGG	GATGTGCCCG
	10201	GCGACGAGTT	GGTGGTTCGCC	GAGCGGCCGG	CCGAAGTCTG	CCCGGGTCCG	GGCGTGGGCC
	10261	ACCGCGGCGG	TGCGGCAGGC	CCGACGAGATC	CCGACGCGAG	CCGAGGCGAC	CGACTTGCGC
60	10321	CCGTAGGCGA	GTGACGCGGC	GACGAGCATC	GGCAGTGACG	CGCCGGAGCC	GGCGAGGACC
	10381	GCGCCGCGCG	GCACACGCAC	TAGGTCCAGG	TGCAGATCGG	CGTGGCCGGC	GGCGGCGCAG
	10441	CCGACGCGCT	TCGGGACGCG	CTCGACGCGT	ACGCCGGGGG	TGTCGGCGGG	CACGACGACC
	10501	ACCGCACCGG	AACCATCCTC	CTGGAGACCG	AAGACGACCA	GGTGGTCCGC	GTAGGCGGGC
	10561	GCAGTCGTCC	AGACCTTGTG	GCGCTCGACG	ACAGCGGTGT	CCCCGTGAG	CCGAACCCGC
	10621	GTCCGCATCG	CCGACAGATC	GCTGCCCGCC	TGCCGCTCAC	TGAAGCCGAC	GGCCGCGAGT

	10681	TTCCTCGCTGG	TCAGCTCCTT	CAGGAAGGTC	GCCCGCTGAC	CGGCGTCGCC	GAGCCGCTGC
	10741	ACGGTCCACG	CGGCCATGCC	CTGCGACGTC	ATGACACTGC	GCAGCGAACT	GCAGAGGCTG
	10801	CCGACGTGTG	CGGTGAACTC	GCCGTTCTCC	CGGCTGCCGA	GTCCCAGACC	GCCGTGCTCG
	10861	GCCGCCACTT	CCGCGCAGAG	CAGGCCGTCG	GCGCCGAGCC	GGACGAGCAG	GTCGCGCGGC
5	10921	AGTTCGCCCG	ACGTGTCCCA	CTCGGCGGCC	CGGTACCCGA	CAAGGTCGGT	CAGCAGCGCG
	10981	TCACGCTCAG	GCATCGACGG	CCCGCAGCCG	GTGGACGAGT	GCGACCATGG	ACTCGACGGT
	11041	ACGGAAGTTC	GCGAGCTGGA	GGTCCGGGCC	GGCGATCGTG	ACGTGCAACG	TCTTCTCCAG
	11101	GTACACGACC	AGTTCCATCC	CGAACAGCGA	CGTGAGGCCG	CCCTCCGCGA	ACAGGTCGCG
	11161	GTCCACGGGC	CAGTCCGACC	TGGTCTTCGT	CTTGAGGAAC	GCGACCAACG	CGTCCGCGAC
10	11221	GGGGTCGTCC	TTGACGGGTG	CGGTCATGAG	AACACCTTCT	CGTATTCTGA	GAAGCCCGCG
	11281	CCGGTCTTCC	GGCCGTGGTG	TCCCTCGCGG	ACCTTGCCCA	GCAGCAGGTC	ACAGGGGCGG
	11341	CTGCGCTCGT	CGCCGGTGCG	TTTGTGCAGC	ACCCACAGCG	CGTCGACGAG	GTTGTCTGAT
	11401	CCGATCAGGT	CCGCGGTGCG	CAGCGGCCCG	GTCCGATGGC	CGAGGCACCC	CGTCATGAGC
	11461	GCGTCGACGT	CCTCGACGGA	CGCGGTGCC	TCCTGCACGA	TCCGCGCCGC	GTCGTTGATC
15	11521	ATCGGGTGGA	GCAGCGGCT	CGTGACGAAG	CCGGGCGCGT	CCCGTCCAGC	GATCGGCTTG
	11581	CGCCGCAGCG	CCGCGACGAG	CTCCCCGGCG	GCGGCCATGG	CCCTCTCACC	GATCCGGGGT
	11641	CCGCGGATCA	CCTCGACCGT	CGGGATCAGG	TACGACGGGT	TCATGAAGTG	CGTGCCGAGC
	11701	AGGTCTCTCG	GCCGGGCCAC	GGAGTCGGCC	AGTTCGTCAA	CCGGGATCGA	CGACGTGTTC
	11761	GTGATGACCG	GGATACCGGG	CGCCGCTGCC	GAGACCGTGG	CGAGTACCTC	CGCTTGACC
20	11821	TCCGGCTCCT	CGACGACGGC	CTCGATCACC	GCGGTGGCCG	TACCGATCGC	GGGACGCGCG
	11881	GACGTGGCCG	TCCGACGAC	ACCGGGTTCG	GCCCTCGGCG	GCCCGGCCAC	GAGTTGTGCC
	11941	GTCCGAGTT	CGGTGGCGAT	CCGCGCCGCG	CCGCGCGTAA	GGATCTCTC	GGAGCTGTCG
	12001	ACGAGTGTCA	CCGGGACGCC	GTGGCGCAGC	GCGAGCGTGG	TGATGCCGGT	GCCATCACT
	12061	CCCGCGCCGA	GCACGATCAG	CTGGTGGTCC	ACGCTGTTTC	CTCCCTCCGG	GGTACCATTG
25	12121	GCAGCGAGTA	CGGGTCGAGG	ACGTCTTCCG	GGGTCGACCC	GATCGCGTCC	TTGCGGCCGA
	12181	GGCCGAGTTC	GTGCGCGAAG	CCGAGCAGCA	CGTCGAACGC	GATGTGGTCC	GCGAACGCGC
	12241	TGCCCCCTCGA	GTGAGGACG	CTCAGGCTGT	CCCGGTGGTC	CGCCGCGGTG	TCCGGTGCCG
	12301	CGCACAGGGC	CGCCAGCGAC	GGGCCGAGCT	CGCGGTCCGG	CAGTTGCTGG	TACTCGCCCT
	12361	CGGCGCGGGC	CTGCCCGGGA	TGGTCGACGC	AGATGAACGC	GTCGTCGAGC	AGGGTCTTCG
30	12421	GCAGTTCGGT	CTTGCCCGGC	TCGTGCGGCG	CGATGGCGTT	CACATGCAGG	TGCGGCAGCC
	12481	GCGGCTCGGC	GGGACGACCC	GGCCCTTTGC	CCGAGGGCAC	CGAGGTGACG	GTGGACAGGA
	12541	CATCCGCGGC	GGCGGCGGCC	TCCGCCGGAT	CGGTACCTT	GACCGGCAGT	CCGAGGAACG
	12601	CGATGCGGTC	CGCGAACGAC	GCCGCGTGCC	CGGGGTCCGT	GTCGCTGACC	AGGATCCGCT
	12661	CGATGGGCAG	GACCTTGCTG	AGCGCGTGCC	CCTGGGTAC	CGCCTGTGCG	CCCGCGCCGA
35	12721	TCAGCGTAG	CGTGGCGGTG	CTCGGACGGG	CCAGCAGCCG	GCTCGCGACG	CGCGCGACCG
	12781	CGCCGTCGCG	CATCGCGGTG	ATCACGCCGT	CGTCGCGGAG	GCGCGTCAGA	CTGCGCGTGT
	12841	CGTCGTCGAG	GCGCGACATC	GTGCCGACGA	TCGTGCGGAG	CCGGAAGCGC	GGATAGTTGT
	12901	GCGGACTGTA	CGAAACCGTC	TTCATGGTCA	CGCCGACACC	GGGACCCCGG	TACGGCATGA
	12961	ACTCGATGAC	GCCGGGAATG	TGCGCGCCGC	GGACGAATCC	GGTACGCGGC	GGCGCCTCGG
40	13021	CGAACTCGCC	GCGGCCGAGC	GCGGCGAACC	CGTCGTGCAG	CTCGCTGATC	AGCCGGTCCA
	13081	TCATCAGCTC	GCGGCCGATC	ACGGAGAGAA	TCCGCTTGAT	GTCACGTTGG	CGCAGGACCC
	13141	TGGTCTGCAT	GTGTCACCTC	CGTTTCTGTT	CCGGAGCTGT	CTTGGTGGTG	CGGCTCGGGG
	13201	CGGCTTCCGT	TCTCATCGCA	GCTCCCTGTC	GATGAGGTGC	AAAATCTCGT	CCGCGGTGCG
	13261	GTCCGCGGAC	AGCACGCCGG	CCGGCGTGGT	CGGGCGGGTC	TCCCGCCGCC	AGCGGTTGAG
45	13321	CAGGGCGTCC	AGCCGGGTTT	CGATCGCGTC	CGCCTGGCGG	GCGCCCGGGT	CGACACCGGC
	13381	AACGAGTGCT	TCCAGCCGGT	CGAGCTGCGC	GAGCACCACG	GTCACCGGGT	CGTCCGGGGA
	13441	CAGCAGTTCA	CCGATGCGGT	CGGCGAGTGC	GCGCGGCGAC	GGGTAGTCGA	AGACGAGCGT
	13501	GGCGGACAGT	CGCAGACCGG	TGCCTCTGTT	GAGGCCGTTG	CGCAGCTGCA	CCGCGATGAG
	13561	CGAGTCCACA	CCGAGTTCCC	GGAACGCGCG	GTCTCTCCGG	ATGTCCTCCG	GGTCGCGGTG
50	13621	GCCCAAGACG	GCCGCTGCCT	TCTGCCGGAC	GAGGGCGAGC	AGGTCCGGTG	GGCGTTCCTG
	13681	CTCGTTGCGG	GCGCTCCGGC	GGGCCGACGG	CTTGGGCCGG	CCACGCAGCA	GCGGGAGGTC
	13741	CGGCGGCGAG	TCGCCCGCCA	CGGCGACGAC	ACTGCCCGTT	CCGGTGTGGA	CGGCGCGGTC
	13801	GTACATGCGC	ATGCCCTGTT	CGGCGGTGAG	CGCGCTCGCC	CCACCCTTGC	GCATACGGCG
	13861	CCGGTCCGGC	TCCGTCAAGT	CCGCGGTGAG	GCCACTCGCC	TGGTCCCACA	GCCCCACGCG
55	13921	GATCGACAGC	CCTGGCAGCC	CTTGTGCACG	CCGCTGTTCC	GCGAGCGCGT	CCGAGAACGC
	13981	GTTCCGCGCC	GCGTAGTTGC	CCTGACCGGG	GGTGCCGAGC	ACACCGGCCG	CCGACGAGTA
	14041	GACGACGAAT	GCGGCGAGGT	CGGTGTGCGG	GGTGAGCCGG	TGCAGGTGCC	AGGCGGCGTC
	14101	GGCCTTGGGT	TTGAGGACGG	TGTCGATGCG	GTCGGGGGTG	AGGTTGTGCA	GCAGGGCGTC
	14161	GTCGAGGGTT	CCGGCGGTGT	GGAAGACGGC	GGTGAGGGGT	TGAGGGATGT	GGGCGAGGGT
60	14221	GGTGGCGAGT	TGGTGGGGGT	CGCCGACGTC	GCAGGGGAGG	TGGGTGCCGG	GGGTGGTGTC
	14281	GGGGGGTGGG	GTGCGGGAGA	GGAGGTAGGT	GTGGGGGTGG	TTCAGGTGGC	GGGCGAGGAT
	14341	CGGCGGAGG	GTGCGGAGC	CGCGGTGAT	GACGACGGCC	CCCTCGGGGT	CCAGCGGCGG
	14401	CGGGACCGTG	AGGACGATCT	TGCCGGTGTG	CTCGCCGCGG	CTCATGGTCC	CCAGCGCCTC
	14461	GCGGACCTGC	CGCATGTCTG	GCACCGTCAC	CGGCGAGGGG	TGCAGCACAC	CGCGCGCGAA

	14521	CAGGCCGAGC	AGCTCCGCGA	TGATCTCCTT	GAGCCGGTGC	GGCCCCGCGT	CCATCAGGTC
	14581	GAACGGTTCG	TGGACGGCGT	GCCGGATGTC	CGTCTTCCCC	ATCTCGATGA	ACCGGCCACC
	14641	CGGCGCGAGC	AGGCCGACGG	ACGCGTCGAG	GAGTTCACCG	GTGAGCGAGT	TGAGCACGAC
5	14701	GTCCAGCGGC	GGGAACGCGT	CGGCGAACGC	GGTGCTGCGG	GAATCGECCA	GATGCGCTCC
	14761	GTCCAGGTCC	ACCAGATGGC	GCTTCGCGGC	GCTGGTGGTC	GCGTACACCT	CCGCGCCCAG
	14821	GTGCCGCGCG	ATCTGCCGGG	CGGCGGAACC	GACACCGCCG	GTGGCCGCGT	GGATCAGGAC
	14881	CTTCTCGCCG	GGGCGCAGCC	CGGCGAGGTC	GACCAGGCCG	TACCACGCGG	TCGCGAACGC
	14941	GGTCATCACG	GACGCCGCCT	GCGGGAACGT	TCAGCCGTCC	GGCATCCGGC	CGAGCATCCG
	15001	GTGGTCGGCG	ATGACCCTGG	GGCCGAAGCC	GGTGCCGACG	AGGCCGAAGA	CGCGCTCGCC
10	15061	CGGTGCCAGA	CCGGAGACGT	CGGCGCCGGT	CTCCAGGACG	ATGCCCGCGG	CCTCGCCGCC
	15121	GAGCACGCCC	TGACCGGGGT	AGGTGCCGAG	CGCGATCAGC	ACATCGCGGA	AGTTGAGGCC
	15181	CGCCGCACGC	ACACCGATCC	GGACCTCGGC	CGGGGCGAGG	GGGCGCCGGG	GCTCCGCCGA
	15241	GTCCGCCGCG	GTGAGGCGGT	CGAGGGTGCC	CGTCCGCGCC	GGCCGGATCA	GCCACGTGTC
	15301	GCTGTCCGCG	ACGGTGAGCG	GCTCCGGCAC	CCGGGTGAGG	CGGGCCGCCCT	CGAACCGGCC
15	15361	GCCGCGCAGC	CGCAGACGCG	TGCGACGCGG	TCGACGCGCG	ATGCGCTGCT	GCTCGGGGGC
	15421	GAGCGTGACG	CCGGACTCGG	TCTCGACGTG	GACGAACCGG	CCGGGCTGCT	CCGGCTGGGC
	15481	GGCGCGCAGC	AGTCCGGCCG	CCGCGCCGGT	GGCGAGGCCG	GCGGTGGTGT	GCACGAGCAG
	15541	ATCCCCGCGG	GAGCCGGTCA	GGGCGGTGAG	CAGCCGGGTG	GTGAGCGCAC	GCGTCTCGGC
20	15601	CACCGGGTCG	TCGCCATCAG	CGGCAAGCAA	CGTGATGACG	TCCACGTCCG	TCGCGGGGAC
	15661	ATCCCTGGGT	GCGGCGACCT	CGATCCAGGT	GAGACGCATC	AGGCCGGTGC	CGACGGGTGG
	15721	GGACAGCGGG	CGGGTGCGGA	CCGTCCGGAT	CTCGGCGACG	AGTTGGCCGG	CGGAGTCGGC
	15781	GACGCGCAGA	CTCAGTCTGT	CGCCCTCAGC	AGTGATCAGC	GCTCGGAGCA	TGGCCGAGCC
	15841	CGTGCGGACG	AACCGGGCCC	CCTTCCAGGC	GAACGGCAGA	CCCGCAGCGC	TGTCGTCCGG
25	15901	CGTGGTGAGG	GCGACGGCGT	GCAGGGCCGC	GTGAGCAGC	GCCGGATGCA	CACCGAAACC
	15961	GTCCGCCCTCG	GCGGCTGTGT	CGTCCGGCAG	CGCCACCTCG	GCATACACGG	TGTCACCATC
	16021	ACGCCAGGCA	GCCCCGAACC	CCTGGAACGC	CGACCCGTAC	TCATAACCGG	CATCCCGCAG
	16081	TTGCTCATAG	AACCCCGAGA	CGTCGACGGC	CACGGCCGTG	ACCGGCGGCC	ACTGCGAGAA
	16141	CGGCTCCACA	CCGACAACAC	CGGGGTGTCT	GGGGGTGTCT	GGGGTCAGGG	TGCCGCTGGC
30	16201	GTGCCGGGTC	CAGCTGCCCG	TGCCCTCGGT	ACGCGCGTGG	ACGGTCACCG	GCCGCCGTCC
	16261	GGCCTCATCA	GCCCCCTTCA	CGGTACCCGA	CACATCCACC	GCTGCGGTCA	CCGGCACC7.C
	16321	AAGGGGGGAT	TCGATGACCA	GCTCGTCCAC	TATCCCGCAA	CCGGTCTCGT	CACCGGCCCG
	16381	GATGACCAGC	TCCACAAACG	CCGTACCCGG	CAGCAGGACC	GTGCCCCGCA	CCGCGTGATC
	16441	AGCCAGCCAG	GGGTGAGTGC	GCAATGAGAT	CCGGCCAGTG	AGAACAACAC	CACCATCGTC
35	16501	GGCGGGCAGC	GCTGTGACAG	CGGCCAGCAT	CGGATGCGCC	GCACCCGTCA	ACCCCGCCGC
	16561	CGACAGATCG	GTGGACCCGG	CCGCTCCAG	CCAGTACCGC	CTGTGCTCGA	ACCGGTACGT
	16621	GGGCGATCC	AGCAGCCGTC	CCGGCACC GG	TTCGACCACC	GTGTCCAGT	CCACTGCCGT
	16681	GCCAGGGTTC	CACGCCCTGC	CCAACGCCGT	CAGCCACCGC	TCCAGCCGCG	CGTCACCGGT
	16741	CCGCAACGAC	GCCACCGTGT	GAGCCTGCTC	CATCGCCGGC	AGCAGCACCG	GATGGGCACT
40	16801	GCACTCCACG	AACACCGACC	CATCCAGCTC	CGCCACCGCC	GCGTCCAACG	CCACCGGACC
	16861	ACGCGATTTC	CGGTACCAGT	ACCCCTCATC	CACCGGCTCC	GTCACCCAGG	CGCTGTCCAC
	16921	GGTCGACCAC	CACGCCACCG	ACGCGGCCCT	CCCTGCCACC	CCCTCCAGTA	CCCTGGCCAG
	16981	TTATCCTCTG	ATGGCTTCCA	CGTGGGGCGT	GTGGGAGGCG	TAGTCCAGCG	CGATCCAGAC
	17041	CACCCGCACG	CCTTCGGCCT	CATACCGCGC	CACCACCTCC	TCCACCGCCG	ACGGGTCCCC
45	17101	CGCCACCACC	GTGGAAGCCG	GGCCGTTACG	CGCCGCGATC	CACACACCCT	CGACCAGACC
	17161	GACCTCACC	GCCGGCAACG	CCACCGAAGC	CATCGCTCCC	CGCCCGGCCA	GTGCGCCGCG
	17221	GATGACCTGA	CTGCGCAATG	CCACCACGCG	GGCGGCGTCC	TCGAGGCTGA	GGGCTCCGGC
	17281	CACGCACGCC	GCCGCGATCT	CGCCCTGGGA	GTGTCCGATC	ACCGCGTCCG	GCACGACCCC
	17341	ATGCGCCTGC	CACAGCGCGG	CCAGGCTCAC	CGCGACCGCC	CAGTGGCCCG	GCTGGACACC
50	17401	CTCCACCCGC	TCCGCCACAT	CCGGCCGCGC	CAACATCTCC	CGCACATCCC	AGCCCCGTGT
	17461	CGGCGACAA	GCCTGAGCGC	ACTCCTCCAT	ACGCGCGGCG	AACACCGCGG	AGTGGGUCAT
	17521	GAGTTCACG	CCCATGCCGA	CCCACTGGGC	GCCCTGGCCG	GGGAAGACGA	ACACCGTACG
	17581	CGCCTGGTCC	ACCGCCACAC	CCGTACCCCG	GGCATCGCCC	AGCAGCACCG	CACGGTGACC
	17641	GAAGACAGCA	CGCTCCCGCA	CCAACCCCTG	CGCGACCGCG	GCCACATCCA	CACCACCCCC
55	17701	GCGCAGATAC	CCCTCCAGCC	GCTCCACCTG	CCCCCGCAGA	CTCACCTCAC	CACGAGCCGA
	17761	CACCGGCAAC	GGCACCACAC	CGTCAACAAC	CGACTCCCCA	CGGACGCGCC	CGGACACACC
	17821	CTCAAGGATC	ACGTGCGCGT	TGCTACCGCT	CACCCCGAAC	GACGACACAC	CCGCATGCGG
	17881	TGCCCCGATC	GACTCGGGCC	ACGGCCTCGC	CTCGGTGAGC	AGCTCCACCG	CACCGGCCGA
	17941	CCAGTCCACA	TGCGACGACG	GCTCGTCCAC	ATGCAGCGTC	TTCGGCGCGA	TCCCGTACCG
60	18001	CATCGCCATG	ACCATCTTGA	TCACACCGGC	GACACCCGCC	GCCGCTGCGG	CATGACCGAT
	18061	GTTCGACTTC	AACGAACCCA	GCAGCAGCGG	AACCTCACGC	TCCTGCCCGT	ACGTCGCCAG
	18121	AATGGCCTGC	GCCTCGATGG	GATCGCCAG	CGTCGTCCCC	GTCCCGTGCG	CCTCCACCAC
	18181	GTCCACATCG	GCGGCGCGCA	GTCCGGCGTT	CACCAACGCC	TGCTGGATGA	CACGCTGCTG
	18241	GGACGGGCGG	TTGGGGGCGG	ACAGCCCGTT	GGAGGCACCG	TCCTGGTTCA	CCGCCGACCC
	18301	GCGGACGACC	GCGAGAACGG	TGTGTCCGTT	GCGCTCGGCG	TCGGAGAGCC	GCTCCAGCAC

	18361	AAGAACGCCG	GCGCCCTCCG	CCCAGCCGGT	GCCGTGGGCG	GCGTCCGCGA	ACGCGCGGCA
	18421	GCGGCCGTCTG	GGGGAGAGTC	CGCCCTGCTG	CTGGAATTCC	ACGAACCCGG	TCGGGGTCCG
	18481	CATGACGGTG	ACACCGCCGA	CCAGCGCCAG	CGAGCACTCC	CCGTGGCGCA	GTGCGTGCCC
	18541	GGCCTGGTGC	AGCGCGACCA	GCGACGACGA	GCACGCCGTG	TCCACCGTGA	ACGCCGGTCC
5	18601	CTGGAGCCCA	TAGAAGTACG	AGATCCGGCC	GGTGAGCACG	CTGGGCTGCA	TGCCGATCGA
	18661	GCCGAACCCG	TCCAGGTCCG	CGCCGACGCC	GTACCCGTAC	GAGAAGGCGC	CCATGAACAC
	18721	GCCGGTGTCTG	CTGCCGCGCA	GTGTGCCCGG	CACGATGCCC	GCGCTCTCGA	ACGCCTCCCA
	18781	TGTCGTTTTCC	AGCAGGATCC	GCTGCTGGGG	GTCCATGGCC	CGTGCCTCAC	GGGGGCTGAT
	18841	GGCGAAGAAC	GCGGCATCGA	AGCCGGCCGG	CTCGGAGAGG	AAGCCGCCGC	GGTCCGTGTC
10	18901	CGATCCGCCG	GTGAGGCCGG	ACGGGTCCCA	GCCACGGTCG	GCCGGGAAGC	GCGTGACCGC
	18961	GTCGCCGCCA	CTGTCCACCA	TGCGCCACAG	GTCGTGCGGC	GAGGTGACGC	CGCCCGGCAG
	19021	TCGGCAGGCC	ATGCCCACGA	TGGCCAGCGG	TTCGTACCGG	GTCGCGCGG	CTGTGGGAAC
	19081	AGCGACCGGT	GCGGCACCAC	CGACCAGAGC	CTCGTCCAAC	CGCGACGCGA	TGGCCCGCGG
	19141	CGTCGGGTAG	TCAAGACAA	GCGTGGCGGG	CAGTCGGACA	CCGGTCGCCG	CGGCGAGTCG
15	19201	GTTCCGCAGT	TCGACGGCGG	TCAGCGAGTC	GATACCCAGT	TCCTTGAAGG	CCGCGTCCCG
	19261	GCGGCGTCCG	GCGGCGTCCG	CGTGCCCGAG	CACCGCCGCC	GCGTTGTCTG	GGACCACTGC
	19321	CAGCAGCGCG	GTGTCCCGCT	CAGCGCCGGA	CATGGTGCCG	AGCCGGTCGG	CGAGCGGAAC
	19381	GGCGGTGGCC	GCCGCCGGGC	GCGATACGGC	GCGGCGCAGA	TCGGCGAAAA	GCGGCGATGT
	19441	GTGCGCGGTG	AGGTCCATCG	TGGCCGCCAC	GGCGAACGCG	GTGCCGGTTC	CGGCCGCGGC
20	19501	TTCCAGCAGG	CGCATGCCCA	CACCGGCCGA	CATGGGGCGG	AAACCGCCGC	GGCGGACACG
	19561	GGTGCGGTTG	GTGCCGCTCA	TGCTGCCGGT	GAGTCCGCTG	TCATCGGCCC	AGAGGCCCAA
	19621	GGCCAGCGAC	AGCGCGGGCA	GTCTTCCGGC	ATGGCGCAGC	GTCGCGAGTC	CGTCGAGGAA
	19681	CCCGTTCCGC	CCCGAGTAGT	TGCCCTGGCC	GCGGCGCGCC	ATGATGCCCG	CGAGGAGCAA
	19741	GTAGAGGACG	AACGAGCGCA	GGTCCGCGTC	CCGGGTCAGC	TCGTGCAGGT	GCCAGGCGCC
25	19801	GTGCGCTTTG	GGGCGCAGTG	TGGTGGCGAG	CCGCTCCGGG	GTGAGTGCCG	TGGTCACGCC
	19861	GTGCTCGAGC	ACGGCTGCCG	TGTGGAAGAC	CGCCGTGAGC	GGCCTGCCGG	CGGCGGCGAG
	19921	CGCGGCGGGC	AGCTGGTCCC	GGTCGGCGAC	GTCACAGCGG	ATGTGGACAC	CGGGAGTGTG
	19981	CGCCGGCGGT	TCGCTGCGCG	ACAGCAACAG	GAGGTGGCGG	GCGCCATGCT	CGGCGACGAG
	20041	ATGCGGGCGG	AGGAGACCTG	CCAGCACACC	CGAGCCGCCG	GTGATGACCA	CCGTGCCGTG
30	20101	CGGGTCGAGC	AGCGGTTCCG	GCGTTTCCGC	GGCGGCCGTG	CGGGTGAACC	CGGCGGCTTC
	20161	GTACCGGCCG	TCGGTGACGC	GGACGTACGG	CTCGGCCAGT	GTCGTGGCGG	CGGCCAGCCC
	20221	CTCGATGGGG	GTGTGCGTGC	CGGTCTCCAC	CAGCACGAAC	CGGCCCGGGT	GCTCGGCTTG
	20281	GGCGGACCGG	ACGAGGCCGG	CGACCGCTCC	TCCGACCGGT	CCCGGCTCGA	TCCGACGAC
	20341	GAGGGTGGTC	TCCGCAGGGC	CGTCTCCGGC	GATCACCCGG	TGCAGCTCGC	CGAGCACGAA
35	20401	CTCGGTGAGC	CGGTACGTCT	CGTCGAGGAC	ATCCGCGCCC	GGTTCGGGGA	GCGCGGAGAC
	20461	GATGTGGACC	GCGTCCGACT	GACCGGGCCC	GGGAGTGGGC	AGTCCGGTCC	AGGAGAGGCC
	20521	GTACAAGGAG	TTCCGTACGA	CGGCGGCGTC	GCCGTGACAG	TTCACCGGTC	GCGCGGTGAG
	20581	CGCGGCGACG	GTCACCACCG	GTTGGCCGAC	CGGGTCCGTC	GCATGCACGG	CAGCGCCGTC
	20641	CGGGCCCTGA	GTGATCGTGA	CGCGCAGCGT	GGTGGCCCGG	GTCGTGTGGA	ACCGCACGCC
40	20701	GCTCCACGAG	AACGGCAGCC	GCACCTCCGC	TTCCTGTTCC	GCGAGCAGCG	GCAGGCAGGT
	20761	GACGTGCAAG	GCCGCGTCTGA	ACAGCGCCGG	GTGGACGCCA	TAGTGCGGCG	TGTCGTCEGC
	20821	CTGTTCCCCG	GCGATCTCCA	CCTCGCGGTA	CAGGGTTTCG	CCGTGCGGCG	AGGCGGTGCG
	20881	CAGTCCCTGG	AACGCTGGGC	CGTAGCTGTA	GCCGGTCTCG	GCCAGCCGCT	CGTAGAACCG
	20941	GCTCACGTCTG	ACGCGTCTCG	CGCCCGGCGG	CGGCCACGCG	GGCGGCGGGA	CCGCCGCGAC
45	21001	GCTTCCGGCC	CGGCCGAGGG	TGCCGCTGGC	GTGCCGGGTC	CAGCTGTCCG	TGCCCTCGGT
	21061	ACGCGCGTGG	ACGGTCACTC	GCCGCCGTCC	GGCCTCATCG	GCCCCCTCGA	CGGTACCCGA
	21121	CACATCCACC	GCGCCGGTCA	CCGGCACACC	GAGCGGGGTC	TCGATGACCA	GTTTATCCAC
	21181	CACCCCGCAA	CCGGTCTCGT	CACCGGCCCG	GATGACCAGC	TCCACAAACG	CCGTACCCGG
	21241	CAGCAGAACC	GTGCCCGCA	CCCGGTGATC	AGCCAGCCAG	GGATGCGTAC	GCAACGAGAT
50	21301	CCGGCCAGTG	AGAACAACAC	CACCACCGTC	GTCGGCGGGC	AGTGCTGTGA	CGGCGGCCAG
	21361	CATCGGATGC	GCCGCCCGCG	TCAGCCCGGC	CGCGGACAGA	TCGGTGGCAC	CGGCCGCTTC
	21421	CAGCCAGTAC	CGCCTGTGCT	CGAACGCGTA	GGTGGGCGAG	TCGAGCAGCC	GTCCCGGCAC
	21481	CGGTTTCGACC	ACCGTGTCCC	AGTCCACTGC	CGTGCCCGAG	GTCCACGCCT	GCGCCAACGC
	21541	CGTCAGCCAC	CGCTCCCAGC	CGCCGTCAAC	GGTCCGCAAC	GACGCCACCG	TGTGAGCCTG
55	21601	TTCCATCGCC	GGCAGCAGCA	CCGGATGGGC	GCTGCACCTC	ACGAACACGG	ACCCGTCCAG
	21661	CTCCGCCACC	GCCGCGTCCA	GCGCAGCGGG	GCGACGACGG	TTCCGGTACC	AGTAGCCCTC
	21721	ATCCACCGGC	TCGGTCACCC	AGGCGCTGTC	CACCGTGGAC	CACCAGGCCA	CCGACCCGGT
	21781	CCCGCCGGAA	ATCCCTCCA	GTACCTCGGC	CAACTCGTCC	TCGATGGCTT	CCACGTGGGG
	21841	CGTGTGGGAG	GCGTAGTCGA	CCGCGATACG	GCGCACTCGC	ACGCCTTCGG	CCTCGTACCG
60	21901	CGTCACCACT	TCTTCCACCG	CGGACGGGTC	CCCCGCCACC	ACAGTCGAAG	ACGGGCGGTT
	21961	ACGCGCCGCG	ATCCACACGC	CCTCGACCAG	GTCCACCTCA	CCGGCCGGCA	ACGCCACCGA
	22021	AGCCATCGCC	CCCCGCCCGC	CCAGCGCCCG	GCGCATCACC	TGGCTGCGCA	AGGCCACAC
	22081	GCGGGCGGCG	TCCTCAAGGC	TGAGGGCTCC	GGCCACACAC	GCCGCCGCGA	TCTCGCCCTG
	22141	GGAGTGTCCG	ACCACCGCGT	CCGGCACGAC	CCCATGCGCC	TGCCACAGCG	CGGCCAGGCT

	22201	CACCGCGACC	GCCCAGCTGG	CCGGCTGGAC	CACCTCCACC	CGTCCGCCA	CATCCGGCCG
	22261	CGCCAACATC	TCCCGCACAT	CCCAGCCCGT	GTGCGGCAAC	AACGCCCCG	CACACTCCTC
	22321	CATACGAGCC	GCGAACACCG	CAGAACACGC	CATCAACTCC	ACACCCATGC	CCACCCACTG
5	22381	AGCACCTGTC	CCGGGAAAGA	CGAACACCGT	ACGCGGCTGA	TCCACCGCCA	CACCCATCAC
	22441	CCGGGCATCG	CCCAACAACA	CCGCACGGTG	ACCGAAGACA	GCACGCTCAC	GCACCAACCC
	22501	CTGCGCGACC	GCGGCCACAT	CCACACCACC	CCCGCGCAGA	TACCCCTCCA	GCCGCTCCAC
	22561	CTGCCCCCGC	AGACTCACCT	CACCTCCGAGC	CGACACCGGC	AACGGCACCA	ACCCATCGAC
	22621	AGCCGACTCC	CCACGCGACG	GCCCAGGAAAC	ACCCTCAAGG	ATCACGTGCG	CGTTCGTACC
10	22681	GCTCACCCCG	AAAGCGGAGA	CACCGGCCCC	GCGCGGACGT	CCCGCGTCGG	GCCACGCCCG
	22741	CGCCTCGGTG	AGCAGTTCCT	CCGCGCCCTC	GGTCCAGTCC	ACATGCGACG	ACGGCTCGTC
	22801	CACATGCAGC	GTCTTCGGCG	CGATGCCATA	CCGCATCGCC	ATGACCATCT	TGATGACACC
	22861	GGGACACCC	GCAGCCGCCT	GCGCATGACC	GATGTTTCGAC	TTCAACGAAC	CCAGCAGCAG
	22921	CGGAACCTCA	CGCTCCTGCC	CGTACGTCGC	CAGAATCGCG	TGCGCTCGA	TGGGATCGCC
15	22981	CAGCGTCGTC	CCCGTCCCGT	GCGCTCCAC	ACGCTCCACG	TGCGCGGGG	CGAGCCJCGC
	23041	CTTGTGGAGG	GCCTGGCGGA	TGACGCGCTG	CTGGGAGGGG	CCGTTGGGTG	CGGAGATGCC
	23101	GTTGAGGCG	CCGTCTTGGT	TGACGCGGGA	GGAGCGGACG	ACCGCGAGGA	CGGTGTGTCC
	23161	GTTGCGCTCG	GCGTCGGAGA	GCTTTTCGAC	GACGAGGACG	CCGGCCCCCT	CGGCGAAACC
	23221	GGTGCCGTCC	GCCGCGTCAG	CGAACGCCTT	GCACCGTCCG	TCCGGCGCGA	CGCCGCCCTG
20	23281	CCGGGAGAAC	TCCACGAAGG	TCTGTGGTGA	TGCCATCACT	GTGACACCAC	CGACCAGCGC
	23341	CAGCGAGCAC	TCCCGGTCC	CGAGCGCTCG	CCCGGCTTGG	TGCAGCCGCA	CCAGCGACGA
	23401	CGAACACGCC	GTGTCGACCG	TGACCGCCGG	ACCCTCCATG	CCGAAGAAGT	ACGACAGCCG
	23461	TCCGGCGAGC	ACCGCGGGCT	GTGTGCTGTA	GGCGCCGAAT	CCGCCAGGT	CCGCGCCCGT
	23521	GCCGTAGCCG	TAGTAGAAGC	CGCCGACGAA	GACGCCGGTG	TCGCTGCCGC	GCAGGGTGTC
25	23581	CGGCACGATG	CCGGCGTGTT	CGAGCGCCTC	CCAGGCGATT	TCGAGGAGGA	TCCGCTGCTG
	23641	CGGGTCGAGT	GCGGTGGCCT	CGCGCGGACT	GATGCCGAAG	AACGCGGCAT	CGAAGTCGGC
	23701	GGCGCCCGCG	AGTGCGCCGG	CCCGCCCGGT	GGCGGACTCG	GCGGCGGCGT	GCAGCGCGGC
	23761	CACGTCCCAG	CCGCGGTCCG	TGGGGAAGTC	GCCGATCGCG	TCCGCGCCGT	CCGCGACGAG
	23821	CTGCCACAGC	TCTTCCGGTG	AGGTGACGCC	GCCCCGCAGT	CGGCAGGCCA	TGCCGACGAC
30	23881	GGCGAGCGGC	TCGTTCCGCG	CGGCGCGCAG	CGCGGTGTTT	TCCCGGCGGA	GCTGCGCGTT
	23941	GTCTTTCGAC	GACGTCCGCA	GCGCCTCGAT	CAGGTGCTTC	TCCGCCATCG	CCTCATCCCT
	24001	TCAGCACGTG	CGCGATGAGC	GCGTCTGCGT	CCATGTCTGTC	GAACAGTTTC	TGTCGCGGCT
	24061	CCGCGGTCTG	GGTGCTCGCG	GGTGCTCTGT	CCGGTGGTTC	ACCGCCGTCC	GGGGTCCCGT
	24121	TGTCTCCCGG	GGTCCCCTTG	ACGTCCGGGG	CCAGGAGGGT	CAGCAGATGA	CGGGTGAGCG
35	24181	GCCCGGCGGC	GGGATAGTCG	AAGACGACGC	TGGCCGCGAG	CGGAATGCCG	AGGGCCTCGG
	24241	AGACCCGGTT	GCGCAGGCCG	AGCGCGGTGA	GCGAGTCGAC	CCCGAGGTCC	TTGAACGCCG
	24301	TGGTGGCCGT	GACCGCCGCC	GCGTCGGTGT	GGCCAGCAG	GGTGGCGGCG	GTGTGCGGGA
	24361	CGACGCCGAG	CAGCACCTGT	TCCCGTTCCT	TGTGGGGCAG	GTCCGGCAGG	CGTTCAGCA
	24421	GGGAGCCGCC	GTCGGTCCGCG	GAGCGCCGGG	TGGGGCGCTG	GATCGGTTCG	CACAGCGGTG
40	24481	ACGGGTTCGCC	GGGCCCCGGT	GGGGCGGTTC	CCACGACCAC	GGCTTCCCCG	GTGGCGCACG
	24541	CGGCGTCGAG	GAGGTCCGTC	AGCCGGTCCG	CCCGGCGGGT	GAACGCCACG	GCCGCGAGGC
	24601	CTTGTGCCCG	GCGCAGGTTC	GCCAGGGCCT	GGAGCGGTCC	GGCCGCTTCG	CCGGACGGAA
	24661	CGGCGAGAAC	GAACCGGGTC	AGGTGAGAGT	CGCGGGTCAG	GCGGTGCAGT	TCCCAGGCCG
	24721	ACTCGGCGGT	GCCGTCCGCG	TGGACGACCG	CGGTACCCGG	GGTTTCCGGC	ACTGTGCCCG
45	24781	GCTCGTACCG	GATCACTTCG	GCGCCGTGTC	CGCCGAGGTG	TCCGGCGAGT	TCCTCCGAAC
	24841	CGCCCGCGAG	GAGGACGGTG	TCGCCGTACG	AGGCCGCGGC	CGTGGTGGGC	GCGGCGGGGA
	24901	CGAGGCGGGG	CGCTTCGAGG	CGCCCGTCCG	CCAGGCGCAG	GTGCGGTTCG	TGAGGCGGGG
	24961	AGAGGGCGGC	GCGCGGCGCG	GGGGTGACCG	TGTCGGTGGT	CTCCACGAGC	ACGAGCCGGC
	25021	CCGTTTCCGC	GGTGTGAGC	AGTGCGGCGA	CGGCACCGGC	GACGGGCCCC	GCCCTCGGCG
50	25081	ACACCACCAG	CGTGGCGCCG	GCGGTCTCTG	GGTCTGTCAG	TGCGGTACGG	ACCTCGTCGG
	25141	GACCGGATAC	CGGGACGACG	ATGACGTCCG	GCGTGGCGTC	GTGCGCGAGG	TCCGTGTACC
	25201	GGCGGGCCGT	GGTGCCGGGT	GCCGCGGGGG	CCCGGACGCC	GGTCCAGGTG	CGCCGGAACA
	25261	GCCGCACGTC	CCCGTCCGGG	CCCGTCCGTC	CGGGGGGCGG	GGTGATGAGC	GAGCCG.TCT
	25321	GAGCCACCGG	CCGTCCAGT	TCGTTCGGGA	GGTGACGCG	GGCGCCGCC	TCCGCTTCGC
55	25381	GCTGGACGAA	GGTGACGCGC	AGTTTCGTGG	GCGCGCTGGT	GTGGACACGG	ACGCGGGTGA
	25441	ACGCGAACGG	CAACCGTACC	CCCGCGTTCT	CGGCGGCCGC	GCCGATGCTG	CCCGCTTGCA
	25501	GCGCGGTGAC	GAGCAGCGCC	GGGTGCAGTG	TGTAGCGGGC	GGCGTCCCTG	GCGAGGGCGC
	25561	CGTCGAGGGC	GACTTCGGCG	CAGACGGTGT	CTCCGTGGCT	CCACGCGGCG	GACATGCCGC
	25621	GGAACCTCGG	GCCGAACTCG	TATCCCGCGT	CGTCGAGTCG	CTGGTAGAAG	GCCGCGACGT
60	25681	CGACCGGTTT	CGCGTGCTCG	GGCGGCCAGG	GCCCCGGCGT	GGTGGCCGGT	TCCGTGGTGG
	25741	CGATGCCGGC	GAAGCCGGAG	GCGTGCGGGG	TCCATGTCCG	GTGCGCGTCC	GTCCGGGCGT
	25801	GGACGCGCAC	GGCACGGCGT	CCGGTGTCGT	CGGGCGCGGC	GACGGTCAAC	CGACCTGGA
	25861	CGGCGCCGGT	GGCGGGCAGG	ACCAGCGGTG	TCTCGACGAC	CAGTTCGTTC	AGCAGGTTCG
	25921	AGCCTGCCTC	GTCGGCGCCG	CGTCCGGCCA	ATTCCAGGAA	GGCGGGTCCG	GGCAGCAGTA
	25981	CGGCGCCGTC	GACGGAGTGA	CCGGCCAGCC	ATGGGTGGGT	GGCCAGCGAG	AACCGGCCCG

	26041	TGAGCAGCAC	CTCGTCGGAG	TCGGGGAGCG	CCACCGACGC	GGCGAGCAGC	GGGTGGTCTGA
	26101	CGGCGTCGAG	TCCGAGGCCG	GAAGCGTCCG	TGCCGGCCGC	GGTCTCGATC	CAGTAGCGCT
	26161	CATGGTGGAA	GGCGTATGTG	GGCAGGTCGT	GTGCCGTGCG	CGTCGCGGGG	ACGACCGCCG
5	26221	CCCAGTCGAC	GGGCACGCCG	GTTGTGTGCG	CCTCGGCCAG	CGCGGTGAGC	AGCCGGTGGGA
	26281	CTCCCCCGCC	GCGGCGGAGC	GTGGCGACGG	TGCGCCGCTC	GATCGCGGGC	AGCAGCACGG
	26341	GGTGCGCGCT	GACCTCGACG	AACACGGTGT	CACCCGGCTC	GCGGGCAGCG	GTCACGGCCG
	26401	TGGCGAAGCC	TACGGGGTGG	CGCATGTTGC	GGAACAGTA	CTCGTCGTCTG	AGCGGCGCGT
	26461	CGATCCAGCG	TTCGTGCGCG	GTGGAGAACC	ACGGGATCTC	GGGCGTGCGC	GAGGTGGTGT
10	26521	CCGCGACGAT	CCGCTGGAGT	TCGTCTGTACA	GCGGGTTCGAC	GAACGGGGTG	TGGTCTGGGC
	26581	AGTCGACGGC	GATGCGGCGC	ACCCAGACGC	CGCGGGCCTC	GTAGTCGGCG	ATCAGCGTTT
	26641	CGACGGCGTC	CGGGCGCCCG	GCGACGGTCG	TGGTGGTGGC	GCCGTTGCGG	CCCGCGACCC
	26701	AGACGCCGTC	GATCCGGGCG	GCATCCGCCT	CGACGTCGGC	GGCCGGGAGC	GCGACCGAGC
	26761	CCATCGCGCC	GCGTCCGGCG	AGTTCGCGCA	GGAGCAGGAG	AACGCTGCGC	AGCGCGACGA
15	26821	GGCGGGCACC	GTCCTCCAGG	GTGAGCGCTC	CGGCGACACA	GGCCGCGGGC	ATCTCGJCCT
	26881	GGGAGTGTC	GATGACGGG	TCCGGGCGTA	CGCCCGCGGC	CTCCCACACG	CGGGCCAGCG
	26941	ACACCATGAC	GGCCAGCAG	ACGGGGTGCA	CGACGTCGAC	GCGGCGGGTC	ACCTCCGGGT
	27001	CGTCGAGCAT	GGCGATGGGG	TCCAGCCCCG	TGTGCGGGAT	CAGCGCGTCG	GCGCATTGGC
	27061	GCATCCTGGC	GGCGAACACC	GGGGAGGCCG	CCATCAGTTC	GACGCCCATG	CCGCGCCACT
20	27121	GCGGTCTTTG	TCCGGGGGAA	ACGAAGACGG	TGCGCGGCTC	GGTGAGCGCC	GTGCCGGTGA
	27181	CGACGTCGTC	GTCGAGCAGC	ACGGCGCGGT	GCGGGAACGT	CGTACGCCTG	GCGAGCAGGC
	27241	CCGCGCGGAT	GGCGCGCGGG	TGCTGGCCGG	GACGGGCGGC	GAGGTGCTCG	CGGAGTCGGC
	27301	GGACCTGGCC	GTCGAGGGCC	GTGGCGGTC	GCGCCGAGAC	GGCGAGTGGT	GTGAGCGGGC
	27361	TGGCGATCAG	CGGCTCACCG	GGCTTCGAGG	CCGACGGCTC	CTCGGCCGGC	GGTCCCCCGG
25	27421	CCGGGTGGGC	TTCCAGCAGG	ACGTGGGCGT	TGGTGCCGCT	GACGCCGAAG	GAGGACACAC
	27481	CGGCGCGCCG	CGGGCGGTCTG	GTCTCGGGCC	AGGGCCGGGC	ATCGGTGAGG	AGTTCGACGG
	27541	CGCCGGCCGT	CCAGTCGACG	TGCGAGGACG	GCGTGTCCAC	GTGCAGGGTG	CGCGGCAGGG
	27601	TGCCGTGCCG	CATGGCGGAG	ACCATCTTGA	TGACACCGGC	GACACCCGCG	GCGGCCTGAG
	27661	GTGGGCGGAT	GTTGGACTTC	AGCGAGCCCC	GCAGACCGG	GGTGTGCGCG	CCCTGCCCGT
30	27721	AGGTGGCCAG	CACCGCCTGT	GCCTCGATGG	GATCGCCAG	CCTGGTGCCG	GTGCCGTGCG
	27781	CCTCCACGGC	GTCCACGTCC	GCCGGGGTGA	GCCCCGCGTT	GGCCAGGGCC	TGCCGGATCA
	27841	CCCCGCTCCTG	CGAGGGCCCC	TTCCGGCGCC	ACAACCCGTT	GGAAGCACCG	TCCTGGTTGA
	27901	CCGCCGAACC	CCGGACAACC	GCCAGCACAC	GGTGGCCGTT	GCGCTCGGCA	TCGGAGAGCC
	27961	TCTCGACGAT	CAGCACACCG	GACCCCTCGG	CGAAACCGGT	GCCGTCAGCC	GCATCCJCGA
35	28021	ACGCCTTGCA	GCGCGCGTCG	GGCGCGAGAC	CCCGCTGCTG	GGAGAACTCG	ACGAAGCCGG
	28081	ACGCGAGGC	CATCACCGTG	ACGCGCGCGA	CGAGGGCGAG	CGAGCATTCC	CCGGAGCGCA
	28141	GTGACTGCCC	GGCCTGGTGC	AGCGCCACCA	GCGACGACGA	ACACGCCGTG	TCGACCGTGA
	28201	CCGCCGGACC	CTCCAGACCG	TAGAAGTACG	ACAGCCGACC	GGACAGCACA	CTGGTCTGGG
	28261	TGCCGGTCTG	GCCGAAACCG	CCCAGGTCGG	TGCCGAGTCC	GTACCCGTCG	GAGAAGGCGC
40	28321	CCATGAACAC	GCCGGTGTCTG	CTTCCGCGCA	GCGACTCCGG	GAGGATCCCG	GCGTGTTCCTA
	28381	GCGCCTCCCA	CGAGGTCTCC	AGGACCAGAG	GCTGTGCGG	GTCCATCGCC	AGCGCCTCAC
	28441	GCGGACTGAT	CCCGAAGAAC	GCCGCGTCGA	AGTCCGCCAC	CCCGGCGAGG	AAGCCACCAT
	28501	GACGACGGT	CGACGTGCCC	GGATGATCCG	GATCGGGATC	GTACGGGCGG	TCCAGCTCCC
	28561	AACCACGGTC	CGTCGGAAC	GCCGTGATCC	CGTCACCACC	CGACTCCAGC	AGCGCCACA
45	28621	AGTCTCTCGG	CGACGCGACC	CCACCCGGCA	GCCGGCAGGC	CATCCCCACG	ATCGCCAACG
	28681	GTCGTCCTG	CCGGACGGCC	GCGGTCTGTTG	TGCGGGTCGG	CGATGCCGTC	CGGCCGACA
	28741	GCGCCGCGGT	GAGCTTCGCC	GCGACGGCGC	GCGGCGTCGG	GAAGTCGAAG	ACCGCGGTGG
	28801	CGGGCAGCCG	TACGCCCGTC	GCCTCGGTGA	AGGCGTTGCG	CAGCCGGATC	GCCATGAGCG
	28861	AGTCGACGCC	GAGTTCCTTG	AACGTGGCGG	TCGCCTCGAC	CCGTGCGGCA	CCGTCTGTGGC
50	28921	CGAGTACGGC	CGCGGTGCAC	TGCCGGACGA	CGGCGAGCAC	GTCCCTTTTCG	GCGTCCGCGG
	28981	CGGAGAGCCG	CGCGATCCCG	TCGGCGAGGG	TGGTGGCGCC	GGCCGCCCCG	CGCCGCGGCT
	29041	CCCGGCGCGG	TGCGCGCAGC	AGGGGCGAGC	TGCCGAGGCC	GGCCGGGTCTG	GCGGCGACCA
	29101	GCGCCGGGTC	CGAGGACCGC	AACGCCGCGT	CGAACAGCGT	CAGTCCGCCT	TCGGCGJTCA
	29161	GCGCCGTAC	GCCGTGCGCG	CGCATGCGGG	CGCCGGTGCC	GACCGTCAGC	CCGCTCTCCG
55	29221	GTTCCTACAG	GCCCCAGGCC	ACGGACAACG	CGGGCAGTCC	GGCTGCCCCG	CGCTGTTCCG
	29281	CGAGCGCGTC	GAGGAACGCG	TTCGCGCGCG	CGTAGTTGCC	CTGTCCGGGG	GTGCCGAGCA
	29341	CACCGGCGGC	CGACGAGTAG	AGGACGAACG	CGGCCAGTTC	CGTGTCTTGG	GTGAGTTCGT
	29401	GCAGGTGCCA	CGCGGCGTCC	ACCTTCGGGC	GCAGCACCGT	CTCGAGCCGG	TCGGGGGTGA
	29461	GCGCGGTGAG	GACGCCGTCG	TCGAGGACGG	CCGCGGTGTG	CACGACGGCC	GTGAGCGGGT
60	29521	GCGCCGGGTC	GATCCCCGCC	AGTACGGAGG	CGAGTTCGTC	CCGGTCTGGCG	ACGTCCGAGG
	29581	CGATCGCCGT	GACCTCGGCG	CCGGGCACGT	CGCTCGCCGT	GCCGCTGCGC	GACAGCATCA
	29641	GCAGCCGGCG	CACGCCGTGG	CGTTCGACGA	GGTGGCGGCT	GATGATGCCG	GCCAGCGTCC
	29701	CGGAGCCACC	GGTGACGAGC	ACGGTGCCGT	CCGGGTGAG	CGCCGGAGCG	TCACCCGCGG
	29761	GGACCGCCGG	GGCCAGACGG	CGGGCGTACA	CCTGGCCGTC	ACGCAGCACC	ACCTGGGGCT
	29821	CATCGAGCGC	GGTGGCCGCT	GCGAGCAGCG	GCTCGGCGGT	GTCCGGGGCG	GCGTCGACGA

	29881	GGACGATCCG	GCCGGGGTGT	TCGGCCTGCG	CGGTCCGCAC	CAGTCCGGCG	GCCGCGGCCG
	29941	ACGCGAGACC	GGGCCCCGTG	TGGACGGCCA	GGACCGCGTC	GGCGTACCGG	TCGTCCGGTGA
	30001	GGAAGCGCTG	CACGGCGGTC	AGGACGCCGG	CGCCAGTTTC	GCGGGTGTCT	TCGAGCGGGG
5	30061	CACCCCGGCC	GCCGTGCGCG	GGGAGGATCA	CCACGTCCGG	GACCGTCGGG	TCGTCCAGGC
	30121	GGCCGGTCTG	CGCGGTCTGT	GGCGGCAGCT	CCGGGAGCTC	GGCCAGCACC	GGGCGCAGCA
	30181	GGCCCGGAAC	GGTCCCGTG	ATCGTCAGGG	GGCGCTGCG	CACGGCGCCG	ATGGTGGCGA
	30241	CGGGCCCCGC	GGTCTCGTCC	GCGAGGTGTA	CGCCGTCAGC	GGTGACGGCG	ACGCGTACCG
	30301	CCGTGGCGCC	GGTGGCGTGG	ACGCGGACGT	CGTCGAACGC	GTACGGAAGG	TGGTCCCCTT
	30361	CCGCGGCGAG	GCGGAGTGCG	GCGCCGAGCA	GCGCCGGGTG	CAGGCCGTAC	CGTCCGGCGT
10	30421	CGGCGAGCTG	TCCGTCCGGC	AGGGCCACTT	CCGCCAGAC	GGCGTCTGTC	TCGGCCAGA
	30481	CGGCGCGCGG	GCGGGGCGAG	GCGGGCCCGT	CCGTGTACCC	GGCTCGGGCC	AGACGGTCGG
	30541	CGATGTCTGT	GGGGTCCACC	GGCCGGGCGG	TGGCGGGCGG	CCACGTTCGAC	GGCATCTCCC
	30601	GCACGGCCGG	GGCCGTCCCG	GGGTCCGGGG	CGAGGATTCC	GTGCGCGTGC	TCGGTCCACT
	30661	CCCCCGCCGC	GTGCCGCGTG	TGCACGTGTA	CCGCGCGGCG	GCCGTCCGCC	CCGGGCGCCC
15	30721	TCACCGTGAC	GGAGAGCGCG	AGCGCACC GG	ACCGCGGCAG	CGTGAGGGGG	CGTGTCCACGG
	30781	TGAACGTGTC	GAGGGCGCCG	CAGCCGGCTT	CGTCGCCCGC	CCGGATCGCC	AGATCCAGGA
	30841	GGCCCGCGGC	GGGCAGCACC	GCGAGGCCGT	GCAGGGAGTG	CGCCAGCGGA	TCGGCGGCGT
	30901	CGACCCGGCC	GGTGAGCACC	AGGTCCGCCG	TGCCGGGCAG	GGTGACCGCC	GCGGTACAGC
	30961	CCGGGTGCGC	GACCGGCGTC	TGTCCGGCCG	GGGCCGCGTC	GCCCGCGGTC	TGGGTGCCGA
20	31021	GCCAGTAGCG	GACCCGCTCG	AACGGGTACG	TCGGCGGGTG	CGAGGCGCGT	GCCGGCGCGG
	31081	GGTCGATGAC	CTTCGGCCAG	TCGACCGTGA	CGCCGTCCGT	GTGCAGCCGG	GCGAGCGCGG
	31141	TCAGGGCGGA	TCGCGGTTCT	TCGTCCGGCG	GCAGCATCGG	GATGCCGTGC	ACGAGTCCGGG
	31201	TCAGGCTCCG	GTCCGGGCGG	ATCTCCAGGA	GCACCGCCCC	GTCGTGCGCG	GCGACCTGTT
	31261	CCCCGAACCG	GACGGTGTGG	CGGACCTGTC	GTACCCAGTA	CTCCGGCGTG	GTGCAGGCGG
25	31321	CGCCCGCGGC	CATCGGGATC	CTCGGCTCGT	GGTACGTACG	GCTCTCCCGG	ACCTTGCGGA
	31381	ACTCCTCGAG	CATCGGCTCC	ATCCGCGCCG	AGTGGAAACG	GTGGCTGGTC	CGCAGGCGGG
	31441	TGAAGCGGCC	GAGCCGGGCG	GCGACGTGCA	GCACCGCCTC	CTCGTCACCG	GAGAGCACGA
	31501	TCGACGCGGG	CCCGTTGACC	GCGGCGATCT	CCACGCCGTC	CCGACGAGCG	CAGACGCGCT
	31561	CCCGTTCCGA	CGCGATCAGC	GCGGCCATCG	CCCCGCCGGA	CGGCAGCGCC	TGCATCAGGC
30	31621	GGGCCCCGTG	GGACACCAGC	CTGCACGCGT	CCTCCAGGGA	CCAGACGCCG	GCGACGTACG
	31681	CGGCGGCCAG	CTCGCCGATC	GAATGGCCCA	CGAAGGCGTC	CGGGCGTACG	CCCCACGCCT
	31741	CGAGCTGTGC	GCCGAGTGCG	ACCTGGAGCG	CGAACACCGC	GGGCTGGGCG	TACCCGGTGT
	31801	CGTGGAGGTC	GAGCCCGGCG	GGCAGTCTGA	GGGCGTCCAG	CACCTCGCGG	CGAGTGCGCG
	31861	CGAAGACGTC	GTAGGCGGCG	GCCAGTCCGT	GCCTCATGCC	GGGACGTTGT	GAGCCCTGTC
35	31921	CGGAGAAGAG	CCACACGAGG	CGGCGGTCCG	GTTCTGCGGC	GCCCGTGACC	GTGTCCGTGC
	31981	CGATCAGCGC	GGCCCGGTGC	GGGAAGGCCG	TGCGGGCGAG	CAGGGCCCGG	GCCACCGCGC
	32041	GCTCGTCCTC	CTCGCCGGTG	GCGAGGTGGG	CGCGCAGGCG	GTGTACCTGT	GCGTCGAGTG
	32101	CCTGCGGGGT	GCGTGCCGAG	AGCAGCAGGG	GCAGCGGTCC	GGTGTCCGGT	GCCGGGGCGG
	32161	GTTCCGGGGG	CGGTCCGGGG	TGGCTTTTCA	GGATGATGTG	AGCGTTGGTG	CCGCTAACGC
40	32221	CGAAGGAGGA	CACCCCGGCG	CGCCGTGGGC	GGTCCGTTTC	GGGCCAGGGG	CGGGCGTCCG
	32281	TGAGGAGTTC	GACGGCGCCG	CGCGTCCAGT	GCACGTGCGA	GGACGCGGTG	TCCACGTGCA
	32341	GGGTCCGCGG	CAGGGTGCCG	TGCCGATGG	CGAGGACCAT	CTTGATGACA	CCGGCGACGC
	32401	CCGCGGCGGC	CTGAGTGTGG	CCGATGTTGG	ACTTCAGCGA	GCCCAGCAGC	ACCGGGGTGT
	32461	CGCGATGCTG	CCCGTAGGTG	GCCAGTACCG	CCTGCGCCTC	GATGGGGTCC	CCCAGCCTGG
45	32521	TCCCGGTGCC	ATGCGCCTCG	ACAGCGTCCA	CATCCGCCCG	GGTGAGCCCG	GCGTTGGCCA
	32581	GCGCCTGCCG	GATCACCCCG	TCCTGCGACG	GCCCCGTTCC	CGCCGACAAC	CCGTTGGAAG
	32641	CACCGTCTTG	GTTGACCGCC	GAACCACGCA	CGACCGCCAG	GACATTGTGG	CCGTGCCGCT
	32701	CGGCGTCCGA	GAGCCTCTCG	ACGATCAGCA	CACCGGATCC	CTCGGCGAAA	CCGTTGCCAT
	32761	CAGCCGCATC	CGCGAACGCC	TTGCAGCGGC	CGTCCGGGGA	GAGGCCCGCG	TGCTGGGAGA
50	32821	AGTCCACGAA	GCCGGACGGC	GAGGCCATCA	CCGTGACGCC	GCCGACCACG	GCGAGCGAGC
	32881	ACTCCCCCGA	GCGCAGCGAC	TGCCCGGCCCT	GGTGCAGCGC	CACCAGCGAC	GACGAACACG
	32941	CCGTGTCCAC	CGTGACCGCC	GGACCCTCCA	AACCGTAGAA	GTACGACAGC	CGACCGGACA
	33001	GCACACTGGT	CTGGGTGCTG	GTGGCACCGA	AACCGCCGCG	GTCCGGTCCA	GTGCCGACCC
	33061	CGTAGAAGTA	GCCGCCCATG	AACACGCCGG	TGTCGCTTCC	GCGCAGCGAC	TCCGGGAGGA
55	33121	TCCCGGCGTG	TTCCAGCGCC	TCCACGAGG	TCTCCAGGAC	CAGACGCTGC	TGCGGGTCCA
	33181	TCGCCAGCGC	CTCACGCGGA	CTGATCCCGA	AGAACGCCGC	GTCGAAGTCC	GCCACCCCGG
	33241	CGAGGAAGCC	ACCATGACGC	ACGGTCGACG	TGCCCCGATG	ATCCGGATCG	GGATCGTACA
	33301	GCCCGTCCAC	GTCCCAACCA	CGGTCCGTTC	GAAACGCCGT	GATCCCGTCA	CCACCCGACT
	33361	CCAGCAGCCG	CCACAAGTCC	TCCGGCGACG	CGACCCACCC	CGGCAGCCGG	CAGGCCATCC
60	33421	CCACGATCGC	CAACGGCTCG	TCCTGCCGGA	CGGCCGCGGT	CGGGGTACGG	CGCGGGGTGG
	33481	TCGCGCGCGC	GCCGGCCAGT	TCTGTCAGGT	GCGCGCGGAG	CGCCTGCGCC	TGCGGGTGGT
	33541	CGAAGACGAG	CGTAGCGGGC	AGCGTCAGGC	CCGTCCGCGT	GGCCAGCCGG	TTGGCGAGTT
	33601	CGACGCGCGT	CAGCGAGTCG	AAGCCCACTT	CCCTGAACGC	GCGCGCGGGT	GCGATGGCGT
	33661	GGGCGTCCCG	GTGGCCGAGC	ACCGCGGCAG	CGCTGGTACG	GACGAGGTCC	AGCATGTCCG

	33721	GCGCGGCCGG	AGGTGCGGAC	GTGCGCCGGA	CGGCCGGCAC	GAGGGTGCGT	AGGACCGGCG
	33781	GGACCCGGTC	GGACGCGGCG	ACGGCGGCGA	GGTCGAGCCG	GATCGGCACG	AGCGCGGGCC
	33841	GGTCGGTGTG	CAGGGCCGCG	TGGAACAGGG	CGAGCCCCTG	TGCGGCCGTC	ATCCGGGTCA
	33901	TGCCGTTGCG	GGCGATGCGG	GCCAGGTGCG	TGGCGGTCAG	CCGCCCGCCC	TGCCGGTGGC
5	33961	CCGCCTCCCA	CAGTCCCCAG	GCGAGCGAGA	CGGCGGGCAG	CCCCTGGTGG	TGCCGGTGGC
	34021	GGGCGAGCGC	GTCGAGGAAC	GCGTTGCCGG	TCGCGTAGTT	GGCCTGACCC	GCGCCGCCGA
	34081	ACGTGGCGGA	TATGGACGAG	TACAGGACGA	ACGCGGCCAG	GTCGAGATCG	CGCGTCAGCT
	34141	CGTGACGGTG	CCAGGCGACG	TCCGCCTTGA	CCCGCAGCAC	GGCGTCCCAC	TGCTCCGGCC
	34201	GCATGGTCGT	CACGGCCGCG	TCGTGACGGA	TCCCGGCCAT	GTGCACGACG	GCGCGCAGCC
10	34261	GCTGGGCGAC	GTCGGGCGAG	ACTCGGGCCA	GCTCGTCGCG	GTCGACGACG	TGGCGGGCCA
	34321	CGTACCGCAC	GCGGTCTGTC	TCCGGCGTGT	CGCGGGGCGG	GCCGTTGCGG	GACACCCAGA
	34381	CGACCTCGGC	GGCCTCGTGC	ACGGTGAGCA	GGTGGTCCAC	GAGGAGGCGG	CCGAGCCCGC
	34441	CGGTGCCGCG	GGTGACGAGG	ACGGTCCCGC	CGGTGACGCG	GGAGGTTCCG	GTGGCCGCGG
	34501	CGACACGGCG	CAGACGGGCC	GCACGCGCTG	TGCCGTCGGC	GACCCGGACG	TGCGGCTCGT
15	34561	CGCCGGCGGC	GAGCCCGGCC	GCTATGGCGG	CGGGCGTGAT	CTCGTCCGCT	TCGATCAGCG
	34621	CGACGCGGCC	GGGATGCTCC	GTCTCCGCGG	TCCGGACCAG	GCCGCCGAGC	GCTTCCGCG
	34681	CGGATCGGCC	GGTACGGGTG	GCCACGATGA	GCCGGGATCG	CGCCGACGCG	GGCTCGGCGA
	34741	GCCAGGTCTG	CACGGTGGTG	AGCAGGTGCG	GGCCGAGCTC	CCGGTCCGCG	GCGCCGGCGG
	34801	AGGTGCCCGG	GTCGCCGGGT	TCCACGGCCA	GGACCACGAC	CGGGGGGTGC	TCGCCGTCGG
20	34861	GCACGTGCGC	GAGGTACGTC	CAGTCGGGGA	CGGGTGACGC	GGGCACGGGC	ACCCAGGCGA
	34921	TCTCGAACAG	CGCCTCGGCA	TCGGGGTTCG	CGGCCCGCAC	GGTCAGGCTG	TCGACGTCAA
	34981	GGACCGGTGA	GCCGTGCTCG	TCCGTGGCGA	CGATGCGGAC	CATGTCGGGG	CCGACGCGTT
	35041	CCAGCAGCAC	GCGCAGCGCG	GTCCGCGCGC	GCGCGTGGAT	CCTCACGCGC	GACCAGGAGA
	35101	ACGCCAGCCG	GCGCCGCTCC	GGGTCCGTGA	AGACCGTCCC	GAGGCGTCCG	AGGCGCGCGT
25	35161	CGAGCAGCAC	GGGGTGACGC	CCGTACCGGG	CGTCGGTGAG	CTGTTCCGCG	AGGCGGACCG
	35221	ACGCGTAGGC	GCGGCCCTCC	CCCGTCCACA	TCGCGGTCAT	GGCCCGGAAC	GCGGGCCCGT
	35281	ACGAGAGCGG	CAGCGCGTCG	TAGAAGCCGG	TCAGGTGCGC	CGGGTCCGCG	TCGGCGGGCG
	35341	GCCAGTCCAC	GGGCTCCGCC	GGACCGCCAG	TGTCCACGCT	CAGCGTCCG	GTCCGACTGA
	35401	GCGCCCAAGG	GCCCCTGCGC	GTACGGCTGT	GCAGACTCAC	CGACCGCCGT	CCGGACACCT
30	35461	CGGTTCCGAC	GGTGGCTGG	ATCTCCGTGT	CGCCGTGCGC	GTCGACACAC	ACCGCGCGGA
	35521	CGATGGTTCAG	CTCCGCGATC	TCCGCGGTGC	CGAGCCGGGC	TCCCGCTTCG	TCGAGCAGTT
	35581	CCACGAGCGC	CGAGCCGGGC	ACGATGACCC	GGCCGTCCAC	CTCGTGGTGC	GCGAGCCAGG
	35641	GCTGACGGCG	TACCGAGACA	CCGCGGTGGC	CAGCGCGCCC	TCGCCGTCGG	GCGAGGTCCA
	35701	CCCACGAGCC	GAGCAGCGGG	TGGCCGGACG	TTCGCCCGCG	TTCGCCGTCG	ATCCAGTACC
35	35761	GGTCACGGCG	GAACGGGTAC	GTGGGCAGCG	GCACCACCCG	ACGCGTCCGC	AACGACGAGG
	35821	TGACGGGCAC	GCCCCGGACC	CAGAGCGCGG	CGAGCGACCG	AGTGAAGCGG	TCCAGGCCCC
	35881	CTACGCTCCG	CCGCAGTGTG	CCGGTGACGA	CCGTATGCGC	ATGCCCGGGC	AGCGGTGCTT
	35941	CCAGTGCGGT	GGTGAGCACG	GGATGCGCGC	TGACCTCGAC	GAACGCGCGG	TATCCGCGGT
	36001	CCGCCAGGTG	CCCGGTGCGG	GCGGCGAACC	GAACGGTGCG	GCGCAGGTTG	TCGTACCAGT
40	36061	AGGCGGCGTC	CGCGGGCCGG	TCCAGCCACG	CCTCGTCCAC	GGTGGAGAAG	AACGGGACGT
	36121	CCGGCGTGCG	CGGAGTGATG	CCGGCGAGAG	CGTCGAGCAG	CGCGCCGCGG	ATCGTTTCGA
	36181	CATGCGCGGT	GTGCGACGCG	TAGTCGACGG	CGATCCGGCG	GGCGCGGGGG	GTGGCGGGCA
	36241	CGAGCTCCTC	CACGGCGTCG	GCCGCACCGG	CGACAACGAT	CGACGCGGGT	CCGTTGACCG
	36301	GCGCGACCTC	CAGGCGCCCC	CCCCACACGG	CGGCGTCGAA	GTCGCGGGGC	GGCACCAGAG
45	36361	CCATGCCGCG	CTGCCCGGCC	AGTTCGGTGG	CGACGAGTCG	GCTGCGGACC	GCGACGACCT
	36421	TCGCGGCGTC	GTCCAGGGTG	AGCACCCCGG	CGACGCAGGC	CGCGGCGACT	TCGCCCTGGG
	36481	AGTGGCCGAC	GACCGCGGCC	GGGGCGACCC	CGTGCGCACG	CCACAGCTCC	GCCAGCGCCA
	36541	CCATCACCGC	GAACGACGCG	GGCTGCACGA	CATCGACCCG	GTCGAACGCG	GGCGCTCCGG
	36601	GCCGCTGGGC	GATGACGTCC	AGCAGGTCCC	ATCCGGTGTG	CGGGGCGAGC	GCCGTGGCGC
50	36661	ACTCGCGGAG	CCGCCGGGCG	AACACGGGCT	CGGTGGCGAG	CAGTTCCGCA	CCCATGCCCG
	36721	CCCACTGGGA	GCCCTGCCCG	GGGAACGCGA	ACACGACACG	TGTGTCGGTG	ACGTGCGCGG
	36781	TTCGCGTCAC	GGCCCCCGGC	ACTTCGGCAC	CACGGGCGAA	CGCCTCCGCC	TCTCGGGCCG
	36841	GCACGACCGC	CCGGTGGCGC	ATGGCCGTCC	GGGTGGTGGC	GAGCGAGTGG	CCGACCGCGG
	36901	CCGCGGCGCC	AGTGAGCGGG	GCCAGCTGTC	CCGCGACGTC	CCGCACTCCC	TCCGGGATCC
55	36961	GGGCCGACAT	CGGCCAGACC	ACGTCTCTCG	GCACCGGCTC	GGCTTCGGGT	GCGGACACGG
	37021	GTGCGGGCGC	GGCGGGGGGC	CCGGGCTCCA	GGACGACATG	GGCGTTGGTG	CCGCTGATGC
	37081	GCAACGACGA	GACACCGGCA	CGCCGGTGGC	GCCCGGTGAC	CGCCACGCGC	TCACTGCGGT
	37141	GCAGCAGCCG	GATGTCGCGG	TCCAGTTCGA	CGTGCCGGGA	CGGCTCGTCG	ACGTGCAGCG
	37201	TGCGCGGCAG	GACGCCGTGC	CGCATCGCCA	TGACCATCTT	GATGACGCGG	GCGACGCCGG
60	37261	CCGCGGCCTG	GGTGTGGCCG	ATGTTCTGACT	TGAGCGAGCC	GATCAGCAGC	GGATGCACGC
	37321	GTTCGCGCCC	GTAGGCCACT	TGCAGGGCCT	GGGCCTCGAC	GGGGTCCGCC	AGACGGGTGC
	37381	CGGTGCCGTC	TGCCTCCACG	GCGTCGACGT	CACCCGGCGC	CAGGCCGGCG	TCGGCGAGCG
	37441	CACGCTGGAT	GACGCGCTGC	TGCGCAGGCC	CGTTCGGGGC	GGACAGCCCG	TTCGACGCGC
	37501	CGTCGGAGTT	GACCGCGGAG	CCGCGCACCA	GCGCCAGCAC	GGGGTGGCCG	TGGCGGGTGG

	37561	CGTCGGAGAG	CCGCTCCAGC	ACCAGGACAC	CGGCGCCCTC	GGCGAAGCTC	GTGCCGTCCG
	37621	CGGTGTCCGC	GAAGGCCTTG	GCACGGCCGT	CGGGGGCGAG	CCCGCGCTGC	CGGGAGAACT
	37681	CGACGAACCC	GGTCGTCTGC	GCCATCACCG	TGACACCGCC	GACCAGGGCG	AGCGAGCACT
5	37741	CCCCCGAGCG	CAGCGACCGC	GCGGCCTGGT	GCAGCGCCAC	CAGCGACGAC	GAACACGCCG
	37801	TGTCGACGGT	GACCGACGGG	CCCTCCAGAC	CGAAGTAGTA	CGAGAGCCGC	CCGGAGAGAA
	37861	CGCTGGTCGG	CGTGCCGGTC	GCCCCGAAAC	CGCCCAGGTC	CACGCCCCGC	CCGTAGCCCT
	37921	GGGTGAACGC	GCCCCATGAAT	ACGCCGGTGT	CGCTGCCGCG	GACGCTTTTCG	GGCAGGATGC
	37981	CCGCTCGTTC	GAACGCCCTC	CACGACGCTT	CGAGGACCAG	ACGCTGCTGC	GGGTCCATCG
10	38041	CCAGCGCCTC	ACGCGGGCTG	ATCCCCAAGA	ACGCGGCGTC	GAAGTCGCTC	GGCCCGGTGA
	38101	GGAAGCCGCC	GTGACGCACG	GAAACCTTGC	CGACCGCGTC	GGGGTTCGGG	TCGTAGAGCG
	38161	CGGCGAGGTC	CCAGCCGCGG	TCGGCGGGGA	ACTCGGTGAT	CGCGTCCCCG	CCGGAGTCGA
	38221	CCAGCCGCCA	CAGGTCTCTC	GGTGACCGCA	CGCCACCGGG	CATCCGGCAC	GCCATGGCCA
	38281	CGATCGCCAG	CGGCTCGTTC	CCCGCCACCG	TCGGTGCGGG	CACGTGCGCC	GCCGGAGCGG
	38341	CAGGGGCCGG	CTCACCCTCG	CGTTCCTCAT	CCAGGCGGGC	GGCGAGCGCG	GCCGGTGTCTG
15	38401	GGTGGTCGAA	GACGGCCGTC	CGGAGAGCC	GTACCCCCGT	CGTCTCGGCG	AGGCTGTTGC
	38461	GCAACCGGAC	ACCGCTGAGC	GAGTCGATGC	CGAGGTCCTT	GAACGCCGTC	GTGGCGGTGA
	38521	TCTCGGAGGC	GTGCGCGTGG	CCGAGCACGG	CGGCCGTGGC	CGCACACACG	ATGGCCAGCA
	38581	GGTCACGATC	GCGGTGCGCG	TCGCGGTGCG	GGTTGTCTCT	CGCACGGGCG	GCGATGCGGC
20	38641	GCTCGGTCCG	CTGCCGGACG	GGCTCGGTGG	GAATCGCCGC	GACCATGAAC	GGCACGTCCG
	38701	CGGCGAGGCT	CGCGTCGATG	AAGTGGGTGC	CCTCGGCCTC	GGTGAGCGGC	CGGAACCCGT
	38761	CGCGCACCCG	CTGCCGGTCG	GCGTCGTCAA	GTGTGTCGGT	GAGGGTGCTG	GTGGTGTGCC
	38821	ACATGCCCCA	GGCGATGGAG	GTGGCGGGT	GGCCGAGGGT	GTGGCGGTGG	GTGGCGAGGG
	38881	CGTCGAGGAA	GGCGTTGGCG	GCGGCGTAGT	TTCTTGTCC	GGGGCTGCCG	AGGACGGCGG
25	38941	CGGCGCTGGA	GTAGAGGACG	AAGTGGGTGA	GGGGTTGGTT	TTGGGTGAGG	TGGTGCAGGT
	39001	GCCAGGCGGC	GTTGGCTTTG	GGGTGGAGGA	CGGTGGTGAG	GCGGTGCGGG	GTGAGGGCGT
	39061	CGAGGATGCC	GTGCTCGAGG	GTGGCGGGCG	TGTGGAAGAC	GGCGGTGAGG	GGTTGGGGGA
	39121	TGTGGGCGAG	GGTGGTGGCG	AGTTGGTGGG	GGTCGCCGAC	GTGCGAGGGG	AGGTGGGTGC
	39181	CGGGGTGGT	GTGCGGGGT	GGGGTGGCGG	AGAGGAGGTA	GGTGTGCGGG	TGGTTCAGGT
30	39241	GGCGGGCGAG	GATGCCGGCG	AGGGTGCCGG	AGCCGCCGGT	GATGATGATG	GCGTGTTCGG
	39301	GGTTGAGGGG	GGTGGTGGTG	GGTGGGGTGG	TGGTGTGGAG	GGGGGTGAGG	TGGGTTCGGT
	39361	GGAGGGTGTG	GTGGGTGAGG	CGGAGGTGGG	GGTGGTCGAG	GGTGGCGAGT	TGGGCCAGGG
	39421	GGAGGGGAGT	GTGGGGGTGG	TCGGTTTTCGA	TGAGGCGGAT	GCGGTGGGGG	TGTTCTGTTCT
	39481	GGGCGGTGCG	GGTGAGGCCG	GTGACGGTGG	CGCCGCGCGG	GTGCGTGGTG	GTGTGGACGA
35	39541	TGAGGGTGTG	GTGCGTGGTG	GTGAGGTGGT	GTGTCAGGGC	GGTCAGGACG	CGGGTGGCGC
	39601	GGGTGTGGGC	GCGGGTGGGT	ATGTCCTCGG	GGTCGTGCGG	GTGGGCGGGC	GATATCAGGA
	39661	CGTGTCCCTC	GGGCAGGTCA	CCGTCTGAGA	CCGCTCGGC	GACCGCGAGC	GATCCCAACC
	39721	GGAGCGGGTT	CGGCCCCGAC	GGGGTGTGCG	CCCGCTCCCT	CAGCACCAGC	GAGTCCACCG
	39781	ACACGACAGG	ACGGCCATCC	GGGTGCGCCA	CGCGCACGGC	GACGCCGGCC	TCCCCCGGGG
40	39841	TGAGGGCGAC	GCGCACCGCG	GCGGCCCCCG	TGGCGTTCAG	GCGCACGCCC	GTCCAGGAGA
	39901	ACGGCAGCTC	GATCCCGCCG	CCCGCGTCCA	GGCGCCCGGC	GTGCAGGGCC	GCGTCGAGCA
	39961	GTGCCGGATG	CACACCGAAA	CCGTCCGCCCT	CGGCGGCCCT	CTCGTCGGGG	AGCGCCACCT
	40021	CGGCATACAC	GGTGTACCA	TCACGCCAGG	CAGCCCGCAA	CCCCTGGAAC	GCCGACCCGT
	40081	ACTCATAACC	GGCATCCCGC	AGTTCGTGAT	AGAACCCCGA	GACGTGACAG	GCCGCGGCCG
45	40141	TGGCCGGCGG	CCACTGCGAG	AACGGCTCAC	CGGAAGCGTT	GGAGGTATCC	GGGGTGTCCG
	40201	CGGTACAGGGT	GCCGCTGGCG	TGCCGGGTCC	AGCTGCCCGT	GCCCTCGGTA	CGCGCGTGGA
	40261	CGGTACCCGG	CCGCCGTCCG	GCCTCATCGG	CCCTTCCAC	GGTCACCGAC	ACATCCACCG
	40321	CTGCGGTGAC	CGGCACCACG	AGCGGGGATT	CGATGACCAG	TTCATCCACC	ACCCCGCAAC
	40381	CGGTCTCGTC	ACCGGCCCGG	ATGACCAGCT	CCACAAACGC	CGTACCCGGC	AGCAGAACCG
50	40441	TGCCCCGCAC	CGCGTGATCA	GCCAGCCAGG	GATGCGTACG	CAATGAGATC	CGGCCGGTGA
	40501	GAACAACACC	ACCACCGTCG	TCGGCGGGCA	GTGCTGTGAC	GGCGGCCAGC	ATCGGATGCG
	40561	CCGCCCCGGT	CAGCCCGGCC	GCGGACAGGT	CGGTGGCACC	GGCCGCCCTC	AGCCAGTACC
	40621	GCCTGTGCTC	GAACGCGTAG	GTGGGCAGAT	CCAGCAGCCG	CCCCGGCACC	GGTTCGACCA
	40681	CCGTGCCCCA	GTCCACCCCC	GCACCCAGAG	TCCACGCCCT	CGCCAACGCC	CCCAGCCACC
55	40741	GCTCCAGGCC	ACCGTCACCA	GTCCGCAACG	ACGCCACCGT	GCGGGCCTGT	TCCATCGCCG
	40801	GCAGCAGCAC	CGATGGGCGA	CTGCACTCCA	CGAACACCGA	CCCCTCCAGC	TCCGCCACCG
	40861	CCGCATCCAG	CGCGACAGGG	CGACGCAGGT	TCCGGTACCA	GTACCCGGCT	TCCACCGGCT
	40921	CGGTACCCCA	GGCGCTGTCC	ACGGTCGACC	ACCACGCCAC	CGACCCGGTC	CCGCCGGAAA
	40981	TTCCCTTCAG	TACCTCAGCG	AGTTCGTCTT	CGATGGCCTC	CACGTGAGGC	GTGTGGGAGG
60	41041	CGTAGTCGAC	CGCGATACGA	CGCACCCGCA	CCCCATCAGC	CTCATACCGC	GCCACCACCT
	41101	CCTCCACCGC	CGACGGGTCC	CCCGCCACCA	CCGTGGAAGC	CGGACCATTA	CGCGCCGCGA
	41161	TCCACACACC	CTCGACCAGA	CCCACCTCAC	CGGCCGGCAA	CGCCACCGAA	GCCATCGCCC
	41221	CCCGCCCGGC	CAGCCGCGCC	GCGATCACCC	GACTGCGCAA	CGCCACCACG	CGGGCCGGCT
	41281	CCTCCAGGCT	GAGGGCTCCG	GCCACACACG	CCGCCGCGAT	CTCCCCCTGC	GAGTGTCCGA
	41341	CCACAGCGTC	CGGCACGACC	CCATGCGCCT	GCCACAGCGC	GGCCAGGCTC	ACCGCGACCG

41401 CCCAGCTGGC CGGCTGGACC ACCTCCACCC GCTCCGCCAC ATCCGACCGC GACAACATCT
 41461 CCCGCACATC CCAGCCCGTG TGCGGCAACA ACGCCCGCGC AACTCCTCC ATACGAGCCG
 41521 CGAACACCGC GGAACGGTCC ATGAGTTCCA CGCCCATGCC CACCCACTGG GCACCCTGCC
 41581 CGGGGAAGAC GAACACCGTA CGCGGCTGAT CCACCGCCAC ACCCATCACC CGGGCATCAC
 5 41641 CCAGCAGCAC CGCACGGTGA CCGAAGACAG CACGCTCACG CACCAACCCC TGCGCGACCG
 41701 CGGCCACATC CACCCACCCC CGCGCAGAT ACCCCTCCAG CCGCTCCACC TGCCCCCGCA
 41761 GACTCACCTC ACCACGAGCC GACACCGGCA ACGGCACCAA CCCATCACCA CCGACTCCA
 41821 CACGCGACGG CCCAGGAACA CCTCCAGGA TCACGTGCGC GTTCGTACCG CTCACCCCGA
 41881 ACGACGACAC ACCCGCATGC GGTGCCGAT CCGACTCGGG CCACGGCCTC GCCTCGGTGA
 10 41941 CGAGTCCAC CGCACCGGCC GACCATCCA CATGCGACGA CGGCTCGTCC ACGTGCAGCG
 42001 TCTTCGGCGC GATCCCATGC CGCATCGCCA TGACCATCTT GATGACACCG CGGACACCCG
 42061 CAGCCGCCTG CGCATGACCG ATGTTCTGACT TGACCGAACC GAGGTAGAGC GCGTGTTCGC
 42121 GGTCTGCCC GTAGGCCGCG AGGACGGCCT GCGCCTCGAT CGGGTCGCCC AGCCGCGTGC
 42181 CCGTGCCGTG CGCCTCCACC ACGTCCACAT CGGCGGCGCG CAGTCCGGCG TTGACCAACG
 15 42241 CCTGCCGGAT CACGCGCTGC TGGGCGACGC CGTTGGGGGC GGACAGTCCG TTGGAGGCAC
 42301 CGTCTGGTT CACCGCCGAG CCGCGACGA CGCGGAGAAC GGTGTGCCCC TTGCGCTCGG
 42361 CGTCGGAGAG CCGCTCCAGC ACGAGAACGC CGACGCCCTC GGCGAAGCCG GTCCCGTCCG
 42421 CCGCGTCGGC GAACGCCTTG CACCGTCCGT CCGGGGAGAG TCCGCGCTGC CGGGAGAACT
 42481 CCACGAGCTC TCGGTGTTC GCCATGACGG TGACACCGCC GACCAGCGCC AGGGAGCACT
 20 42541 CCGCGGCCCG CAGTGCCTGT GCGCCTGGT GCAGGGCGAC CAGCGACGAC GAGCAGCCG
 42601 TGTCGACCGT GACCGCCGGG CCTGAAGTC CGTACACGTA CGAGAGGCGC CCGGACAGGA
 42661 CGTCTGTCTG CGTCGCCGTG ACACCGAGCC CGCCAGGTC CCGGCCGACG CCGTAGCCCT
 42721 GGTGAACGC GCCATGAAC ACGCCGGTGT CGCTCTCCCG GAGCCTGTCC GGCACGATCG
 42781 CCGCGTTCTC GAACGCCTCC CAGGAGGTCT CCAGGATCAG GCGCTGTGG GGTCCATCG
 25 42841 CCAGCGCCTC GTTCGACTG ATGCCAAGA ACGCGGCGTC GAACCCGGCG CCGGCCAGGA
 42901 ATCCGCCGTG GCGTGTCTG GAGCGGCCG CCGCGTCCGG GTCCGGGTCC TACAGCGCGT
 42961 CGACGTCCCA GCGCCGGTCG GTGGGGAAC CCGGTGATCGC CTCGGTACCG GCGGCGACGA
 43021 GCGGCCACAG GTCCTCCGGG GAGGCGACCC CGCCGGGCG TCGGCACGCC ATGCCGACGA
 43081 TCGCGACGGG GTCGCCGGAG CCGAGGGTCT GGGCGGTGCG GGTGCGCGT GTCGCGGAGC
 30 43141 CCGCGAGGTG GCGCGGAAC GCACGCGGAG TGGGGTGGTC GAACGCGGTT GCGCGGGCA
 43201 CCGCAGACC CGTCCGCGCG GCGACGGTGT TGGTGAATC GACGGTGGTG AGCGAGTCGA
 43261 GGCCGTTCTC GCGGAACGTG CCGTCCGGGG AGCAGTGTCC GGCGCCCGG AGGCCAGGA
 43321 CCGTGCGGAC GCTGTGCGCG ACCAGGTCGA GCAGTACGTC CTCCCGGCC GCACGGGCCG
 43381 CCGCGAGGCG GTTCGCCAC TCCTGTTCCG TGGCGTCCGG CTCGGCCGGT CCGGTCAGTG
 35 43441 CCGTGAGGAT CCGCGGCGTG GCGCCGCCCA TCGTCGCGG CCGCGCCCG GCGGAACCGG
 43501 TCCGGGCCAC GATGTACGAG CCGCGCCCG CGATGGCCTT CTCGATCAGG TCGCGGCTCA
 43561 GCGCCGGCCG TTCGATGCCG GGCAGCGCGC GGACGGTGAC GGTGGGGAGT CCTCCGCGG
 43621 CCGTGCGCG GGTGTGGCG TCGCGCCGG CCGGGCCGTC GAGCAGGACG TGCACGAGCG
 43681 CGCCGGGGTT CGCGGCTTCC TCGGCTGCGG TGGTCACTG GGTGAGGCCG GTCTCGTCCG
 40 43741 GGAGCAGGCC GCGGACGGTG TCGGCGTCTT CCGCGGTGAC CAGGACCGGC GCGTCCGGGC
 43801 CGATCGGAGG CGGCACGGTG AGGACCATCT TGCCGGTGTG CCGGGCGTGG CTCATCCACG
 43861 CGAACGCGTC CCGCGCACGG CCGATGTCCC ACGGTGTCAC CCGCAGCGGG CACAGCTCAC
 43921 CGCGGTGAA CAGGTGAGG AGCAGTTCGA GGATCTCCCG CAGGCGCGG GATCCACGT
 43981 CCGCCAGGTC GAACGGCTGC TGGGCGGCGT GCGGATGTC GGTCTTGCCC ATCTCGACGA
 45 44041 ACCGCGCGCC CCGTGCGAGC AGGCCGATGG ACGCGTCGAG GAGTTCACCG GTGAGCGAGT
 44101 TGAGCACGAC GTCGACCGGC GGAAGGTGT CCGCGAACGC GCGCTGCGG GAGTTCGCCA
 44161 CATGGTGGT GTCGAAGCCG TCGGCGTGCA GCAGGTGTTG TTTGGCGGGA CTGGCGGTGG
 44221 CGTACACCTC GCGCGCGAGG TGGCGGGCGA TCCGGGTGCG CGCCATGCCG ACACCGCCCG
 44281 TCGCGCGTG GACCAAGACC TTCTGGCCGG GTCGAGCTC GCCCGGTCG ACGAGGCCGT
 50 44341 ACCAGCGGT GCGGAACACG ATGGGACCG ACGCGCGAT GGGGAACGAC CATCCCCGTG
 44401 GGATCCGTGC GACCAAGCCG CCGTCCGCGA CCACGCTGCG CCGGAACGCG TCCTGCACGA
 44461 GACCGAACAC GCGGTGCGCG GGGGCCAGGT CGTCGACGCC GGTCCGACT TCGGTACGA
 44521 TGCCCGCGGC CTCCCCGCC ATCTCGCCCT CGCCCGGGTA GGTGCCGAGC GCGATCAGCA
 44581 CGTCGCGGAA GTTCAGCCCC GCGGCGCGGA CGTCGATGCG GACCTCGCCG GCGGCCAGGG
 55 44641 GCGCGCGGG ACGTCGAGCG GCGCGACGAG GAGGTGCGCG AGCGTTCGG AGGCGGGCGG
 44701 GCGCAGCGC CACTGGCGCG GTGCGGAGG GGGTGGTGT CCGCGCTACC AGCCGGGCA
 44761 CGTAGGCCAC GCCGGCCCCG AGCGCGATCT GGGGTTCGCC GAGCGAGGCC GCGGCGGGGA
 44821 CGAGTCTGTC ATCGCCGTCC GTGTCCACCA GCACGAACGA TCGGGTTCG GCGGCTGGC
 44881 GCGCAGCGC CTCGTCCCAG AGCCGGGCCCT GGTCCGCGTC CCGGATCTCG GCGGGCCGA
 60 44941 CGCCACCGC GCGGCGGGTG ACGACCGTCC GCGGGGTGA CCGGGTGCCG GGCAGGTGCG
 45001 GCGCTCCCA GACCAAGTTC CACAGCGTGG CCTCGCCACT GCGGTGGCG ACCAGATGGG
 45061 CCGCAGCCC CGCGAGCCG GCGCGCTGGA CTTTGGCCGA CCGGTGCGG GGTATCGTGG
 45121 TGACGTGCA GATCTGTCG GGCACCTTGA AGTAGGCGAG CCGGCGCGG CACTCGGCGA
 45181 GGATCGCCTC GCGGGGACG CCGGGGCCGT CGGAAACGAC GTAGAGCAGG GGTATGTGCG

	45241	CGAGGACGGG	GTGCGGGCGG	CCCGCCGCGG	CGGCGTCCCG	GACACCGGCC	ACCTCCTGGG
	45301	CGACGGTCTC	GATCTCCCGG	GGGTGGATGT	TCTCCCGGCC	GCGGATGATC	AGCTCCTTGA
	45361	CCCGGCCGGT	GATCGTCACG	TGTCCGGTCT	CGGCCTGACG	TGCGAGGTCC	CCGGTGCGGT
5	45421	ACCAGCCGTC	CACGAGCACC	TGGGCGGTCT	CCTCCGGCTG	GGCGTGGTAG	CCGAGCATGA
	45481	GGTCGGCCC	GCTCGCCAC	AGCTCGCCCT	CCTCGCCGGG	TGCCACGTCT	GCGCCGGACA
	45541	CCGGGTCGAC	GAACCGCAGC	GACAGGCCCG	GCACGGGCAG	CCCGCACGAG	CCGGGAACCC
	45601	GCGCATCCTC	CAGGGTGTTG	GCGGTGAGCG	AGCCGGTCGT	CTCGGTGCAG	CCGTACGTGT
	45661	CGAGCAGGGG	CACGCCGAAC	GTCGCCTCGA	AATCCCTGGT	GAGCGACGCC	GGCGAGGTGG
10	45721	ATCCGGCGAC	CAGCGCCACG	CGCAGCGCGC	GAGCCCGCGG	CTCGCCGGAC	ACGGCGCCGA
	45781	GGAGGTAGCG	GTACATCGTC	GGCAGCCCGA	CGAGCACGGT	GCTGGAGTGT	TCGGCCAGGG
	45841	CGTCGAGGAC	GTCACGCGCG	ACGAAGCCGC	CCAGGATACG	GGCGGACCGG	CCGACCGTGA
	45901	GGACGGCGAG	CAGGCAGAGG	TGGTGGCCGA	GGCTGTGGAA	CAGCGGGGCG	GGCCAGAGCA
	45961	GTTCTGCTC	CTCGGTGAGC	CGCCAGGACG	GCACGTGCGA	GTGCATCGCG	GACCAACGGC
15	46021	CGCTGCGCTG	TGCGGAAACC	ACGCCCTTGG	GACGGCCGGT	GGTGCCGGAG	GTTGTAGACA
	46081	TCCAGGCGGG	TTCGTCCAGG	CCGAGGTCGT	CGCGGGGCGG	GCACGGCGGC	TCGGTCCCGG
	46141	CGAGGTCCCT	GTAGGAGACG	CAGTCCGGTG	CCCGGCGCCC	GACGAGCACG	ACGGTGGCGT
	46201	CGGTGCCGGT	GCGGCGCACC	TGGTCGAGGT	GGGTTCGCTC	GGTGACCAGC	ACGGTCGCGC
	46261	CGGAGTCCCT	CAGGAAGTGG	GCGAGTTCGG	CGTCGGCGGC	GTCCGGGTTG	AGCGGGACCG
20	46321	CGACGGCGGG	GGCGCGGGCG	GCGGCGAGGT	AGACCTCGAT	GGTCTCGATC	CGGTTGCCGA
	46381	CGAGCATGCG	GACCCGGTCG	CCCGCGTCCA	GCGCGGACGC	GGCGAGGTGT	CCGGCGAGCC
	46441	GGCCGGCCCC	GAGCCGGAGT	TGCGTGACG	TCACGGCGCG	TTGGGAATCC	GTGTAGGCGA
	46501	TCCGGTCGCC	GCGTCGCTCG	GCATGGATGC	GGAGCAATTC	GTGCAACGGC	CGGATTGGTT
	46561	CCACACGCGC	CATGGAAACA	CCTTCTCTCT	GACCAACCGC	ACAACAGCAC	GGAACCGGCC
25	46621	ACGAGTAGAC	GCCGGCGACG	CTAGCAGCGT	TTTCCGGACC	GCCACCCCTT	GAAGATCCCC
	46681	CTACCGTGCG	CGGCCTCCCC	GGACGCTCAT	CTAGGGGGTT	GCACGCATAC	CGCCGTGCGT
	46741	AATTGCCTTC	CTGATGACCG	ATGCCGGACG	CCAGGGAAGG	GTGGAGGCGT	TGTCCATATC
	46801	TGTCACGGCG	CCGTATTGCC	GCTTCGAGAA	GACCGGATCA	CCGGACCTCG	AGGGTGACGA
	46861	GACGGTGCTC	GGCCTGATCG	AGCACGGCAC	CGGCCACACC	GACGTGTGCG	TGGTGGACGG
30	46921	TGCTCCCCGG	ACCGCCGTGC	ACACCACGAC	CCGTGACGAC	GAGGCGTTCA	CCGAGGTCTG
	46981	GCACGCACAG	CGCCCTGTCT	AGTCCGGCAT	GGACAACGGC	ATCGCCTGGG	CCCGCACCGA
	47041	CGCGTACCTG	TTCGGTGTCG	TGCGCACCGG	CGAGAGCGGC	AGGTACGCCG	ATGCCACCGC
	47101	GGCCCTCTAC	ACGAACGTCT	TCCAGTCTAC	CCGGTCGCTG	GGGTATCCCC	TGCTCGCCCC
	47161	GACCTGGAAC	TACGTCAGCG	GTATCAACAC	GACGAACGCG	GACGGTCTGG	AGGTGTACCG
35	47221	GGACTTCTGC	GTGGGCCGCG	CCCAGGCGCT	CGACGAGGGC	GGGATCGACC	GGCCACCAT
	47281	GCCCGCGGCC	ACCGGTATCG	GCGCCACCGG	GGGCGGCATC	ACCTGCGTGT	TCCTCGCCGC
	47341	CCGGGGCGGA	GTCCGGATCA	ACATCGAGAA	CCCCGCCGTC	CTACGGCCCC	ACCACTACCC
	47401	GACGACGTAC	GGTCCGCGGC	CCCCGGTCTT	CGCACGGGCC	ACCTGGCTGG	GCCCCCGCA
	47461	GGGGGGCCGG	CTGTTTCATCT	CCGCGACGGC	CGGCATCCTC	GGACACCGAA	CGGTGCACCA
40	47521	CGGTGATGTC	ACCGGCCAGT	GCGAGGTCGC	CCTCGACAAC	ATGGCCCGGG	TCATCGGCGC
	47581	GGAGAACCAG	CGGCGCCACG	GCGTCCAGCG	GCGGCGACGTC	CTCGCCGACG	TGACCAACCT
	47641	CAAGGTCTAC	GTCCGCCGCC	GCGAGGATCT	CGATACGGTC	CGCCGGGTCT	GCGCCGACG
	47701	CCTGTGAGC	ACCGCGGCCG	TCGCCCTTTT	GCACACCGAC	ATAGCCCGCG	AGGATCTGCT
	47761	CGTCGAAATC	GAAGGCATGG	TGGCGTGACA	ATACCCGGTA	AAAGGCCCGC	GACGCTGCGC
45	47821	CTCGGCGGAT	CCGCGAAGAG	AAAGAAGAGC	GTCACCGCAC	AGCGCGGCAG	CCCGGTCTTT
	47881	TCGTCTTCTG	CACAGCGGCG	GATCTGGTTT	CTCCAGCAAT	TGGACCCGGA	GAGCAACGCC
	47941	TATAATCTCC	CGCTCGTGCA	ACGCTCTGCG	GGTCTATTGG	ACGCGCCGGC	CCTGGAGCGT
	48001	GCGCTGGCGC	TCGTCTGTCG	GCGCCACGAG	CGGTTGCGGA	CGGTGTTCGA	CACCGCCGAC
	48061	GGCGAGCCCC	TCCAGCGGGT	GCTTCCCGCC	CCGGAACACC	TCCTGCGCCA	GCGCGGGGCG
50	48121	GGCAGCGAGG	AGGACGCCGC	CCGGCTCGTC	CGCGACGAGA	TCGCCGCGCC	GTTGACCTC
	48181	GCCACCGGGC	CGTTGATCAG	GGCCCTGCTG	ATCCGCCTCG	GTGACGACGA	CCACGTTCTC
	48241	GCGGTGACCG	TGCACCATGT	CGCCGGCGAC	GGCTGGTCTG	TCGGGCTCCT	CCAACATGAA
	48301	CTCGCAGCCC	ACTACACGGC	GCTGCGCGAC	ACTGCCCGCC	CTGCCGAATG	GCCGCCGTTG
	48361	CCGGTGACGT	ACGCCGACTT	GCGCCGCTGG	GAGCGGCGCG	AACTACCGCG	GCGCCGACTG
55	48421	GACAGGCGTC	TGGCCTACTG	GCGCGAGCAA	CTCCGGGGCG	CCCCGGCGCG	GCTCGCCCTC
	48481	CCCACCGACC	CTCCCCGCCC	GCCGGTCGCC	GACGCGGACG	CGGGCATGGC	CGAGTGGCGG
	48541	CCGCCGGCCG	CGCTGGCCAC	CGCGGTCTCT	ACGCTCGCGC	GCGACTCCGG	TGCGTCCGTG
	48601	TTCATGACCC	TGCTGGCGGC	CTTCCAAGCG	GTCCTCGCCC	GGCAGGCGGG	CACGCGGGAC
	48661	GTGCTGGTTC	GCACGCCCCG	GGCGAACCCT	ACGCGGGGCG	CGTACGAGGG	CCTGATCGGC
	48721	ATCTTCGTCA	ACACGCTCGC	GCTGCGCGGC	GACCTCTCGG	GCGATCCGTC	GTTCCGGGAA
60	48781	CTCTTCGACC	GCTGCCGGGC	CACGACACCG	GACGCGTTTC	CCACGCGCGA	CTGCCGTTTC
	48841	GAGAACGTCA	TGGAATCTCT	CGCACCAGAA	CGCGACCTGT	CGGTCAACCC	GGTCTGTCAG
	48901	GTGCTGTTGC	AGGTGCTGCG	GCGCGACGCG	GCGACGGCCG	CGCTGCCCGG	CATCGCGGCC
	48961	GAACCGTTCC	GCACCGGACG	CTGGTTCACC	CGCTTCGACC	TGGAATTCCA	TGTGTACGAG
	49021	GAGCCGGGTG	GCGCGCTGAC	CGCGAACTG	CTCTACAGCC	GTGCGCTGTT	CGACGAGCCA

49081 CGGATCACGG GGTGCTGGA GGAGTTCACG GCGGTGCTTC AGGCGGTCAC CGCCGACCCG
 49141 GACGTACGGC TGTCGCGGCT GCCGGCCGGC GACGCGACGG CGGCAGCGCC CGTGGTGCCC
 49201 TCGAACGACA CGGCGCGGGA CTTGCCGTC GACACGCTGC CGGGCCTGCT GGCCCGGTAC
 49261 GCCGCACGCA CCCCCGGCGC CGTGCGCGTC ACCGACCCGC ACATCTCCCT CACCTACGCG
 5 49321 CAGCTGGACC GGCGGGCGAA CCGCTCGCG CACCTGCTCC GCGCGCGCGG CACCGCCACC
 49381 GGCGACCTGG TCGGGATCTG CGCGGATCGC GGCGCCGACC TGATCGTGGG CATCGTGGGG
 49441 ATCCTCAAGG CGGGCGCCGC TTATGTGCCG CTGGACCCCG AACATCCTCC GGAGCGCACG
 49501 GCGTTCGTGC TGGCCGACGC GCAGCTGACC ACGGTGGTGG CGCACGAGGT CTACCGTTCC
 49561 CGGTCCCCG ATGTGCCGCA CGTGGTGGCG TTGGACGACC CGGAGCTGGA CCGCGACCCG
 10 49621 GACGACACGG CGCCGGACGT CGAGCTGGAC CGGGACAGCC TCGCCTACGC GATCTACACG
 49681 TCCGGGTCGA CCGGCAGGCC GAAGGCCGTG CTCATGCCGG GTGTACGCGC CGTCAACCTG
 49741 CTGCTCTGGC AGGAGCGCAC GATGGCCGC GAGCCGGCCA GCCGACCGT CCAGTTCGTG
 49801 ACGCCACGT TCGACTACTC GGTGCAGGAG ATCTTTTCCG CGCTGCTGGG CGGCACGCTC
 49861 GTCATCCCGC CGGACGAGGT GCGGTTTCGAC CCGCCGGGAC TCGCCCGGTG GATGGACGAA
 15 49921 CAGGCGATTA CCCGATCTA CGCGCCGACG GCCGTACTGC GCGCGCTGAT CGAGCAGCTC
 49981 GATCCGACA GCGACAGCT CGCCGCCCTG CGGCACCTGT GCCAGGCGG CGAGGCGCTG
 50041 ATCCTCGACG CGCGGTTGCG CGAGCTGTGC CGGCACCGGC CCCACCTGCG CGTGACAAT
 50101 CACTACGGTC CGGCCGAAAG CCAGCTCATC ACCGGGTACA CGCTGCCCGC CGACCCCGAC
 50161 GCGTGGCCCG CCACCGCACC GATCGGCCCG CCGATCGACA ACACCCGCAT CCATCTGCTC
 20 50221 GACGAGGCGA TGCGGCCGGT TCCGGACGGT ATGCCGGGGC AGCTCTGCGT CGCCGGCGTC
 50281 GGCTCGCCC GTGGGTACCT GGCCCGTCCC GAGCTGACCG CCGAGCGCTG GGTGCCGGGA
 50341 GATCGGTCG GCGAGGAGCG CATGTACCTC ACCGGCGACC TGGCCCGCCG CGCGCCCGAC
 50401 GGCGACCTGG AATTCTCGG CCGGATCGAC GACAGGTCA AGATCCGCGG CATCCGCGTC
 50461 GAACCGGGTG AGATCGAGAG CCTGCTCGCC GAGGACGCCC GCGTCACGCA GGCGCGGGTG
 25 50521 TCCGTGCGCG AGGACCGGCG GGGCGAGAAG TTCTGGCCG CGTACGTCGT ACCGGTGGCC
 50581 GGCCGGCACG GCGACGACTT CGCCGCGTCG CTGCGCGCGG GACTGGCCGC CCGGCTGCCC
 50641 GCCGCGCTCG TGCCCTCCGC CGTCGTCTCG GTGGAGCGAC TGCCGAGGAC CACGAGCGGC
 50701 AAGGTGGACC GGCGCGCGCT GCCCGACCCG GAGCCGGGCC CGGCGTCGAC CGGGGGGGTT
 50761 ACGCCCCGCA CCGATGCCGA GCGGACGGTG TGCCGATCT TCCAGGAGGT GCTCGACGTC
 30 50821 CCGCGGGTCG GTGCCGACGA CGACTTTTC ACGCTCGCGG GGCATCCCT CTCTGCCACC
 50881 CCGGTCTGCT CCCGATCCG CGCCGAGCTG GGTGCCGATG TCCGCTGCG TACGCTCTTC
 50941 GACGGGCGGA CGCCCGCCGC GCTCGCCGT GCGGCGGACG AGGCCGGCCC GGCCGCCCTG
 51001 CCCCCGATCG CGCCCTCCGC GGAGAACGGG CCGGCCCCCC TCACCGCGGC ACAGGAACAG
 51061 ATGCTGCACT CGCACGGCTC GCTGCTCGCC GCGCCCTCCT ACACGGTCGC CCCGTACGGG
 35 51121 TTCCGGCTGC GCGGGCCACT CGACCGCGAA CGGCTCGACG CGGCACTGAC CCGGATCGCC
 51181 CGCGGCCACG AGCCGCTGCG GACCGGGTTC CGCGATCGGG AACAGTCTGT CCGCGCGCCC
 51241 GCTCCGGTGC GCGCCGAGGT GGTTCGGTG CCGGTGCGCG ACGTCGACGC CGCGGTCCGG
 51301 GTCGCCACC GGGAGCTGAC CCGGCCGTTT GACCTCGTGA ACGGGTCGTT GCTGCGTGCC
 51361 GTGCTGCTGC CGCTGGGCGC CGAGGATCAC GTGCTGCTGC TGATGCTGCA CCACCTCGCC
 40 51421 GGTGACGGAT GGTCTTTCGA CCTCTGGTC CGGGAGTTGT CGGGGACGCA ACCGGACCTT
 51481 CCGGTGTCTT ACACGGACGT GGCCCGGTGG GAACGGAGTC CGGCCGTGAT CGCGGCCAGG
 51541 GAGAACGACC GGGCCTACTG GCGCCGCGCG CTGGGGGGCG CCACCGCGCC GGAGCTGCC
 51601 GCGGTCCGGC CCGCGGGGGC ACCGACCGGG CGGGCGTTCC TGTGGACGCT CAAGGACACC
 51661 GCGTCTCTGG CGGCACGCCG GGTGCGGGAC GCCCACGACG CGACGTTGCA CGAAACCGTG
 45 51721 CTCGGCGCCT TCGCCCTGGT CGTGGCGGAG ACCGCCGACA CCGACGACGT GCTCGTCGCG
 51781 ACGCCGTTCC CGGACCGGGG GTACGCGGG ACCGACCACC TCATCGGCTT CTTCGGAAG
 51841 GTCCTCGCGC TGCGCCTCGA CCTCGGCGGC ACGCCGTCTG TCCCGAGGT GCTGCGCCGG
 51901 CTGCACACCG CGATGGTGGG CGCGCACGCC CACGAGCGG TGCCCTACTC CGCGCTGCGC
 51961 GCCGAGGACC CCGCGCTGCC GCGGCCCCCC GTGTCGTTCC AGCTCATCAG CGCGCTCAGC
 50 52021 CGGGAACGTC GGCTGCCCCG CATGCACACC GAGCCGTTCC CCGTCGTCGC CGAGACCGTC
 52081 GACGAGATGA CCGGCGAACT GTCGATCAAC CTCTTCGACG ACGGTGCGAC CGTCTCCGGC
 52141 GCGGTGGTCC ACGATGCCGC GCTGCTCGAC CGTGCCACCG TCGACGATTT GCTCACC CGG
 52201 GTGGAGGCGA CGCTGCGTGC CGCCGCGGGC GACCTCACC TACGCGTCAC CGGTTACGTG
 52261 GAAAGCGAGT AGCCATGCCC GAGCAGGACA AGACAGTCGA GTACCTTCGC TGGGCGACCG
 55 52321 CGGAACCTCA GAAGACCGCT CGGGAACCTC CCGCGCACAG CGAGCCGTTG GCGATCGTGG
 52381 GGTGCGCTG CCGGCTGCCC GCGGGGTCG CGTCGCGGA GGACCTTGCG GAGTTGCTGG
 52441 AGTCCGGTGG CGACGGCATC ACCGCGTTCC CCACGAGCCG GGGCTGGGAG ACCACCGCCG
 52501 ACGGTGCGCG CGGCTTCCTC ACCGGGGCGG CCGGCTTCGA CGCGGCGTTC TTCGGCATCA
 52561 GCCCGCGUGA GCGCTGGCG ATGGACCGGC AGCAGCGCCT GGCCCTGGAG ACCTCGTGGG
 60 52621 AGGCGTTTCA GCACGCGGGC ATCGATCCGC AGACGCTGCG GGGCAGTGAC ACGGGGGTGT
 52681 TCCTCGGCGC GTTCTTCCAG GGTACGGCA TCGGCGCCGA CTTCGACGGT TACGGACCA
 52741 CGAGCATTTA CACGAGCGTG CTCTCCGGCC GCCTCGCGTA CTTCAGGCT CTGGAGGCTC
 52801 CGGCGGTAC GGTGACACG CGTGTTCGT GCTGCTGGT GCGCTGTCAC GGCCTGGG
 52861 AGTCGCTGCG CTCCGGCGAA TGCTCGCTCG CCCTGGTCCG CGGCGTCACG GTGATGGCCT

	52921	CGCCGGCGGG	GTTCGCGGAC	TTCTCCGAGC	AGGGCGGCCT	GGCCCCGAC	GCGCGCTGUA
	52981	AGGCCTTCGC	GGAAGCGGCT	GACGGCACC	GTTTCGCCGA	GGGGTCCGGC	GTCTGATCG
	53041	TCGAGAAGCT	CTCCGACGCC	GAGCGCAACG	GCCACCGCGT	GCTGGCGGTC	GTCCGGGGTT
5	53101	CCGCCGTCAA	CCAGGACGGT	GCCTCCAACG	GGCTGTCCGC	GCCGAACGGG	CCGTGCGAGG
	53161	AGCGGGTGAT	CCGGCAGGCC	CTGGCCAACG	CCGGACTCAC	CCCGGCGGAC	GTGGACGCCG
	53221	TCGAGGCCCA	CGGCACCGGC	ACCAGGTGG	GCGACCCCAT	CGAGGCACAG	GCCGTGCTGG
	53281	CCACCTACGG	GCAGGGGCGC	GACACCCCTG	TGCTGCTGGG	CTCGCTGAAG	TCCAACATCG
	53341	GCCACACCCA	GGCCGCGCGC	GGCGTCGCCG	GTGTCATCAA	GATGGTCCTC	GCCATGCGGC
10	53401	ACGGCACCCT	GCCCCGCACC	CTGCACGTGG	ACACGCCGTC	CTCGCACGTC	GA CTGGACGG
	53461	CCGGCGCCGT	CGAACTCCTC	ACCGACGCC	GGCCCTGGCC	CGAAACCGAC	CGCCACGGC
	53521	GCGCCGGTGT	CTCCTCCTTC	GGCGTCAGCG	GCACCAACGC	CCACATCATC	CTCGAAAGCC
	53581	ACCCCGGACC	GGCCCCGAA	CCCGCCCCG	CACCCGACAC	CGGACCGCTG	CCGTGCTGC
	53641	TCTCGGCCCG	CACCCGCGAG	GCATCGAGT	CACAGGTACA	CCGCTGCGC	CGGTTCTCTG
15	53701	ACGACAACCC	CGGCGCGGAC	CGGGTCGCCG	TCGCGCAGAC	ACTCGCCCGG	GCGACCCAGT
	53761	TCGAGCACCG	CGCCGTGCTG	CTCGGCGACA	CGCTCATCAC	CGTGAGCCCG	AACGCCGGCC
	53821	GCGGACCGGT	GGTCTTCGTC	TACTCGGGGC	AAAGCACGCT	GCACCCGCAC	ACCGGGCGGC
	53881	AACTCGCGTC	CACCTACCCC	GTGTTTCGCC	AAGCGTGGCG	CGAGGCCCTC	GACCACCTCG
	53941	ACCCACCCA	GGGCCCGGCC	ACGCACTTCG	CCCACCAGAC	CGCGCTCACC	GCGTCTCTGC
20	54001	GGTCTGGGG	CATCACCCTG	CACGCGGTCA	TCGGCCACTC	CCTCGGTGAG	ATCACCGCCG
	54061	CGGACGCCG	CGGTGTCCTG	TCCCTGAGGG	ACGCGGGCGC	GCTCTCTACC	ACCCGCACCC
	54121	GCCTGATGGA	CCAAGTGGCG	TCGGGCGGCG	CGATGGTCAC	CGTCTGTACC	AGCGAGGAAA
	54181	AGGCAAGCCA	GGTGCTGCGG	CCGGGCGTGG	AGATCGCCGC	CGTCAACGGC	CCCCACTCCC
	54241	TCGTGCTGTC	CGGGGACGAG	GAAGCCGTAC	TCGAAGCCGC	CCGGCAGCTC	GGCATCCACC
25	54301	ACCGCCTGCC	GACCCGCCAC	GCCGGCCACT	CCGAGCGCAT	GCAGCCACTC	GTGCCCCCCC
	54361	TCCTCGACGT	CGCCCGGACC	CTGACGTACC	ACCAGCCCCA	CACCGCCATC	CCCGGCGACC
	54421	CCACACCGC	CGAATACTGG	GCGCACCAGG	TCCGCGACCA	AGTACGTTTC	CAGGCGCACA
	54481	CCGAGCAGTA	CCCGGGCGCG	ACGTTCTCTG	AGATCGGCC	CAACCAGGAC	CTCTCGCCGC
	54541	TCGTGACGG	CGTTGCCGCC	CAGACCGGTA	CGCCCGACGA	GGTGCGGGCG	CTGCACACCG
30	54601	CGCTCGCGCA	GCTCCACGTC	CGCGGCGTCG	CGATCGACTG	GACGCTCGTC	CTCGGCGGGG
	54661	ACCGCGCGCC	CGTCACGCTG	CCCACGTATC	CGTTCCAGCA	CAAGGACTAC	TGGTTCGGGC
	54721	CCACCTCCCG	GGCCGATGTG	ACCGGCGCGG	GGCAGGAGCA	GGTGGCGCAC	CCGTGCTCG
	54781	GCGCCCGGGT	CGCGCTGCCC	GGCACGGGCG	GAGTCGTCCT	GACCGGCCGC	CTGTGCTGG
	54841	CCTCCATCC	GTGGCTCGGC	GAGCACGCGG	TCGACGGCAC	CGTGCTCCTG	CCCGGCGCGG
35	54901	CCTTCCTCGA	ACTCGCGGCG	CGCGCCGGCG	ACGAGGTCGG	CTGCGACCTG	CTGCGACGAA
	54961	TCGTGATCGA	GACGCGGCTC	GTGCTGCCCC	CGACCGGGCG	TGTGGCGGTC	TCCGTGCGAG
	55021	TCGCCGAACC	CGACGACACG	GGGCGGCGGG	CGGTACCCGT	CCACGCGCGG	GCCGACGGCT
	55081	CGGGCCTGTG	GACCCGACAC	GCCGGCGGAT	TCCTCGGCAC	GGCACCGGCA	CCGGCCACGG
	55141	CCACGGACCC	GGCACCCCTG	CCGCCCGCGG	AAGCCGGACC	GGTCGACGTC	GCCGACGTCT
40	55201	ACGACCGGTT	CGAGGACATC	GGGTACTCCT	ACGGACCGGG	CTTCCGGGGG	CTGCGGGGCG
	55261	CCTGGCGCGC	CGGCGACACC	GTGTACGCCG	AGGTGCGGCT	CCCCGACGAG	CAGAGCGCCG
	55321	ACGCCGCCCG	TTTCACGCTG	CACCCCGCGC	TGCTCGACGC	CGCGTCCAG	GCCGGCGCGC
	55381	TGGCCGCGCT	CGACGCACCC	GGCGGGGCGG	CCCGACTGCC	GTTCTCGTTC	CAGGACGTCC
	55441	GCATCCACGC	GGCCGGGGCG	ACGCGGCTGC	GGGTACGGT	CGGCCGCGAC	GGCGAGCGCA
45	55501	GCACCGTCCG	CATGACCGGC	CCGGACGGGC	AGCTGGTGGC	CGTGGTGGT	GCCGTGCTGT
	55561	CGCGCCCGTA	CGCGGAAGGC	TCCGGTGACG	GCCTGCTGCG	CCCGGTCTGG	ACCGAGCTGC
	55621	CGATGCCCGT	CCCGTCCGCG	GACGATCCCG	GCGTGGAGGT	CCTCGGCGCC	GACCCGGGCG
	55681	ACGGCGACGT	TCCGGCGGCC	ACCCGGGAGC	TGACCGCCCG	CGTCTCGGAT	GCGCTCCAGC
	55741	GCCACCTGTC	CGCCGCGGAG	GACACCACCT	TGGTGGTACG	GACCGGCACC	GGCCCGGCCC
50	55801	CTGCCGCGCG	CGCGGGTCTG	GTCCGCTCGG	CGCAGGCGGA	GAACCCCGGC	CGCGTCGTGC
	55861	TCGTGAGGCG	GTCCCGGAC	ACCTCGGTGG	AGCTGCTCGC	CGCGTGCGCC	GCGCTGGACG
	55921	AACCGCAGCT	GGCCGTCCCG	GACGGCGTGC	TCTTCGCGCC	GCGGCTGCTC	CGGATGTCCG
	55981	ACCCCGCGCA	CGGCCCGCTG	TCCCTGCCCG	ACGGCGACTG	GCTGCTCACC	CGGTCCGCCT
55	56041	CCGGCACGTT	GCACGACGTC	GCGCTCATAG	CCGACGACAC	GCCCCGGCGG	GCGCTCGAAG
	56101	CCGGCGAGGT	CCGCATCGAC	GTCCGCGCGG	CCGGACTGAA	CTTCCGCGAT	GTGCTGATCG
	56161	CGCTCGGGAC	GTACACCGGG	GCCACGGCCA	TGGGCGGGCA	GGCCGCGGGC	GTGCTGGTGG
	56221	AGACCGGGCC	CGGCGTGGAC	GACCTGTCCC	CCGGCGACCG	GGTGTTCGGC	CTGACCCGGG
	56281	GCGGCATCGG	CCCGACGGCC	GTCACCGACC	GGCGCTGGCT	GGCCCGGATC	CCCGACGGCT
	56341	GGAGCTTCAC	CACGGCGGGC	TCCGTCCCCA	TCGTGTTCGC	GACCGCGTGG	TACGGCCTGG
60	56401	TCGACCTCGG	CACACTGCGC	GCCGGCGAGA	AGGTCTCTCG	CCACGCGGCC	ACCGCGGGTG
	56461	TCGGCATGGC	CGCCGCACAG	ATCGCCCGCC	ACCTGGGCGC	CGAGCTCTAC	CGACCGGCA
	56521	GTACCGGCAA	GCAGCACGTC	CTGCGCGCGG	CCGGGCTGCC	CGACACGCAC	ATCGCCGACT
	56581	CTCGGACGAC	CGCGTTCCGG	ACCGCTTTCC	CGCGCATGGA	CGTCTCTCTG	AACGCGCTGA
	56641	CCGGCGAGTT	CATCGACGCG	TCGCTCGACC	TGCTGGACGC	CGACGGCCGG	TTGCTCGAGA
	56701	TGGGCGGCAC	CGAGCTGCGC	GACCCGGCCG	CGATCGTCCC	GCGCTACCTG	CCGTTCGACC

56761 TGCTGGACGC GGGCGCCGAC CGCATCGGCG AGATCCTGGG CGAACTGCTC CGGCTGTTCC
 56821 ACGCGGGCGC GCTGGAGCCG CTGCCGTTCC GTGCCTGGGA CGTCCGGCAG GCACGCGACG
 56881 CGCTCGGCTG GATGAGCCG GCCCCCACA TCGGCAAGAA CGTCCTGACG CTGCCCCGGC
 56941 CGCTCGACCC GGAGGGCGCC GTCGTCTCA CCGGCGGCTC CGGCACGCTC GCCGGCATCC
 57001 TCGCCCGCCA CCTGCGCGAA CGGCATGTCT ACCTGCTGTC CCGGACGGCA CCGCCCGAGG
 57061 GGACGCCCCG CGTCCACCTG CCCTGCGACG TCGGTGACCG GGACCAGCTG GCGGCGGCC
 57121 TGGAGCGGGT GGACCGGCCG ATCACCGCCG TGGTGCACCT CGCCGGTGCG CTGGACGACG
 57181 GCACCGTCGC GTCGCTCACC CCGAGCGTT TCGACACGGT GCTGCGCCCG AAGGCCGACG
 57241 GCGCCTGGTA CCTGCACGAG CTGACGAAGG AGCAGGACCT CGCCGCGTTC GTGCTCTACT
 57301 CGCTCGGCGC CGGCGTGCTC GGCAACGCGG GCCAGGGCAA CTACGTCCG CCGAACGCGT
 57361 TCCTCGACGC GTCGCGCGAG CTGCGCCACG GTTCCGGGCT GCCGGCTTC TCGATCGCCT
 57421 GGGGGCTCTG GGAGGACGTG AGCGGGCTCA CCGGCGGCTC CGGCGAAGCC GACCGGGACC
 57481 GGATGCGGCG CAGCGGTTTC CGGGCCATCA CCGCGCAACA GGGCATGCAC CTGTACGAGG
 57541 CGGCCGGCCG CACCGGAAGT CCGGTGGTGG TCGCGGCGGC GCTCGACGAC GCGCCGGACG
 57601 TGCCGCTGCT GCGCGGCCTG CCGCGGACGA CCGTCCGGCG GGCCGCCGTC CGGGAGTGTT
 57661 CGTCCGCGCA CCGGCTCGCC GCGCTGACCG GCGACGAGCT CGCCGAAGCG CTGCTGACGC
 57721 TCGTCCGGGA GAGCACCGCC GCGGTGCTCG GCCACGTGGG TGGCGAGGAC ATCCCCGCGA
 57781 CCGCGGCGTT CAAGGACCTC GGCATCGACT CGCTACCGC GGTCCAGCTG CGCAACGCC
 57841 TCACCGAGGC GACCGGTGTG CCGGTGAACG CCACGGCGGT CTTGCACTTC CCGACCCGCG
 57901 ACGTGCTCGC CGGGAAGCTC GCGGACGAAC TGACCGGCAC CCGCGCGCCC GTCGTGCCCC
 57961 GGACCGCGGC CACGGCCGGT GCGCACGACG AGCCGCTGGC GATCGTGGGA ATGGCCGTC
 58021 GGCTGCCCGG CCGGGTCCGC TCACCCGAGG AGCTGTGGCA CCTCGTGGCA TCCGGCACCG
 58081 AGCCCATCAC GGAGTTCCCG ACCCGACCGG GCTGGGACGT CGACCGCATC TACGACCCGG
 58141 ACCCCGACGC GATCGGCAAG ACCTTCGTCC GGCACGGTGG CTTCTCACC GCGCGACAG
 58201 GCTTCGACGC GCGGTTCTTC GGCATCAGCC CGCGCGAGGC CCTCGCGATG GACCCGACG
 58261 AGCGGGTGCT CCTGGAGACG TCGTGGGAGG CGTTGGAAG CCGCGGCATC ACCCCGGACT
 58321 CGACCCGCGG CAGCGACACC GCGGTGTTCC TCGGCGCCTT CTCCTACGGT TACGGCACCG
 58381 GTGCGGACAC CGACGGCTTC GCGCGACCG GCTCGCAGAC CAGTGTGCTC TCGGGCCGGC
 58441 TGTCGTACTT CTACGGTCTG GAGGGTCCGG CCGTCACGGT CGACACGGCG TGTTCTGCTG
 58501 CGCTGGTGGC GCTGCACGAG GTCGGGACGT CGCTGCGCTC CGGCGAATGC TCGCTGCCCC
 58561 TGGTCGGCGG CGTCACGGTG ATGGCGTCTC CCGGCGGCTT CGTGGAGTTC TCCCGGACG
 58621 GCGGCCTCGC GCGGACGGC CCGGCGAAGG CGTTCGGCGG GGGTGGCGAC GGCACGAGCT
 58681 TCGCCGAGGG TGCCGGTG TGATCGTCG AGAGGCTCTC CGACGCCSAA CGCAACGGTC
 58741 ACACCGTCCT GGCGGTGCTC CGTGGTTCCG CCGTCAACCA GGATGGTGCC TCCAACGGGC
 58801 TGTCGGCGCC GAACGGGCGG TCGCAGGAGC GGGTGATCCG GCAGGCCCTG GCCAACCGCG
 58861 GGTCACCCC GCGGACGTC GACGCGTCCG AGGCCACGG CACCGGCCAC AGGCTGGGCG
 58921 ACCCCATCGA GGCACAGCG GTACTGGCCA CCTACGGACA GGAGCGCGCC ACCCCCTGCG
 58981 TGCTGGGCTC GCTGAAGTCC AACATCGGCC ACGCCAGGC CCGCTCCGGC GTCGCCGGCA
 59041 TCATCAAGAT GGTGCAGGCC CTCGGGACG GGGAGCTGCC GCCGACGCTG CACGCCGACG
 59101 AGCCGTCGCC GCACGTCGAC TGGACGGCCG GCGCCGTCGA ACTGCTGACG TCGGCCGCGC
 59161 CGTGGCCCGA GACCGACCGG CCACGGCGTG CCGCCGTCTC CTCGTTCCGG GTGAGCGGCA
 59221 CCAACGCCCA CGTCATCCTG GAGGCGGGAC CCGTAACGGA GACGCCCGCG GCATCGCCTT
 59281 CCGGTGACCT TCCCTGCTG GTCTCGGAC GATCACCAGG AGCTACCGGA AGCAGATCC
 59341 GCCGACTGCG CGCCTACCTG GACACACCCC CCGACGTCGA CCGGGTGCCG GTGGACAGA
 59401 CGCTGGCCCG GCGCACACAC TTCGCCACC GCGCCGTGCT GCTCGGTGAC ACCGTCATCA
 59461 CCACACCCCC CCGGACCGG CCGACGAAC TCGTCTTCGT CTACTCCGGC CAGGGCACCC
 59521 AGCATCCCGC GATGGGCGAG CAGCTCGCCG CCGCCCATCC CGTGTTCGCC GACGCCTGGC
 59581 ATGAAGCGCT CCGCCGCCTT GACAACCCCG ACCCCACGA CCCCACGCAC AGCCAGCATG
 59641 TGCTCTTCGC CCACAGGGC GCGTTACCG CCCTCCTGCG GTCCTGGGGC ATACCCCCG
 59701 ACGCGGTCAT CGGCCACTCG CTGGGCGAGA TCACCGCGGC GCACGCGCGC GGCATCCTGT
 59761 CGCTGGACGA CGCGTGACCC CTGATACCA CGCGCGCCCG CCTCATGCAC ACGCTCCCGC
 59821 CACCCGGTGC CATGGTCACC GTACTGACCA GCGAAGAGAA GGCACGCCAG GCGTTGCGGC
 59881 CCGGCGTGGA GATCGCCGCC GTCAACGGGC CCCACTCCAT CGTGCTGTCC GGGGACGAGG
 59941 ACGCCGTGCT CACCGTCGCC GGGCAGCTCG GCATCCACCA CCGCCTGCCC GCCCGGACG
 60001 CCGGGCACTC CGCGCACATG GAGCCCGTGG CCGCCGAGCT GCTCGCCACC ACCCGCGGGC
 60061 TCCGCTACCA CCCTCCCCAC ACCTCCATT CCGAACGCC GTGCTGTTCC ACGCCACGC CGGACGCGC
 60121 CCGAGCAGGT CCGCAAGCCC GTGCTGTTCC ACGCCACGC CGAGCAGTAC CCGGACGCCG
 60181 TGTTCTGTGA GATCGGCCCC GCGGAGGACC TCTCCCGCT CGTCGACGGG ATCCCGCTGC
 60241 AGAACGGCAC CGCGGACGAG GTGCACGCGC TGCACACCGC GCTCGCGCAC CTCTACGCGC
 60301 GCGGTGCCAC GCTCGACTGG CCGCGCATCC TCGGGGCTGG GTCACGGCAC GACGCGGATG
 60361 TGCCCGCGTA CGCGTTCCAA CCGCGGCACT ACTGGATCGA GTCGGCACGC CCGGCCGCAT
 60421 CCGACGCGGG CCACCCCGTG CTGGGCTCCG GTATCGCCCT CGCCGGGTG CCGGGCCGGG
 60481 TGTTACGGG TTCCGTGCCG ACCGGTGCGG ACCGCGCGGT GTTCGTGCGC GAGTGGCGC
 60541 TGGCCGCGC GGACGCGGTC GACTGCGCCA CCGTCGAGCG GCTCGACATC GCCTCCGTGC

	60601	CCGGCCGGCC	GGGCCATGGC	CGGACGACCG	TACAGACCTG	GGTCGACGAG	CCGGCCGGACG
	60661	ACGGCCGGCG	CCGGTTACCC	GTGCACACCC	GCACCGGCGA	CGCCCCGTGG	ACGCTGCACG
	60721	CCGAGGGGGT	GCTGCGCCCC	CATGGCACGG	CCCTGCCCGA	TGCGGCCGAC	GCCGAGGGC
5	60781	CCCCACCGGG	CGCGGTGCCC	GCGGACGGGC	TGCCGGGTGT	GTGGCGCCGG	GGGGACCAGG
	60841	TCTTCGCCGA	GGCCGAGGTG	GACGGACCGG	ACGGTTTCGT	GGTGCACCCC	GACCTGCTCG
	60901	ACGCGGTCTT	CTCCGCGGTC	GGCGACGGAA	GCCGCCAGCC	GGCCGGATGG	CGCGACCTGA
	60961	CGGTGCACGC	GTCCGACGCC	ACCGTACTGC	GCGCCTGCCT	CACCCGGCGC	ACCGACGGAG
	61021	CCATGGGATT	CGCCGCCTTC	GACGGCGCCG	GCCTGCCGGT	ACTCACCGCG	GAGGCGGTGA
	61081	CGCTGCGGGA	GGTGGCGTCA	CCGTCCGGCT	CCGAGGAGTC	GGACGGCCTG	GACCGGTTGG
10	61141	AGTGGCTCGC	GGTCGCCGAG	GCGGTCTACG	ACGGTGACCT	GCCCCGAGGA	CATGTCTTGA
	61201	TCACCGCCGC	CCACCCCGAC	GACCCCGAGG	ACATACCCAC	CCGCGCCCAC	ACCCGCGCCA
	61261	CCCGCGTCCT	GACCGCCCTG	CAACACCACC	TCACCACCAC	CGACCACACC	CTCATCGTCC
	61321	ACACCACCAC	CGACCCCGCC	GGCGCCACCG	TCACCGGCCT	CACCCGCACC	GCCCAGAACG
	61381	AACACCCCA	CCGCATCCGC	CTCATCGAAA	CCGACCACCC	CCACACCCCC	CTCCCCCTGG
15	61441	CCCAACTCGC	CACCTTCGAC	CACCCCGACC	TCCGCCTCAC	CCACCACACC	CTCCACCACC
	61501	CCCACCTCAC	CCCCCTCCAC	ACCACCACCC	CACCCACCAC	CACCCCCCTC	AACCCCGAAC
	61561	ACGCCATCAT	CATCACCGGC	GGCTCCGGCA	CCCTCGCCGG	CATCCTCGCC	CGCCACCTGA
	61621	ACCACCCCCA	CACCTACCTC	CTCTCCCGCA	CCCCACCCCC	CGACGCCACC	CCCGGCACCC
	61681	ACCTCCCTTG	CGACGTGCGC	GACCCCGACC	AACTCGCCAC	CACCTCACC	CACATCCCCC
20	61741	AACCCCTCAC	CGCCATCTTC	CACACCGCCG	CCACCTCGA	CGACGGCATC	CTCCACGCCC
	61801	TCACCCCGA	CGCCTCACC	ACCGTCTCTC	ACCCCAAAGC	CAACGCCCGC	TGGCACCTGC
	61861	ACCACCTCAC	CCAAAACCAA	CCCCTACCCC	ACTTCGTCTC	CTACTCGAGC	GCCGCCJCCG
	61921	TCCTCGGCAG	CCCCGGACAA	GGAAACTACG	CCGCCGCCAA	CGCCTTCCTC	GACCCGCTCG
	61981	CCACCCACCG	CCACACCCCTC	GGCCAACCCG	CCACCTCCAT	CGCCTGGGGC	ATGTGGCACA
25	62041	CCACCAGCAC	CCTCACCGBA	CAACTCGACG	ACGCCGACCG	GGACCGCATC	CGCCCGGGCG
	62101	GTTTCCTCCC	GATCACGGAC	GACGAGGGCA	TGCGCCTCTA	CGAGGCGGCC	GTCGGCTCCG
	62161	GCGAGGACTT	CGTCATGGCC	GCCGCGATGG	ACCCGGCACA	GCCGATGACC	GGCTCCGTAC
	62221	CGCCCATCCT	GAGCGGCCTG	CGCAGGAGCG	CGCGGCGCGT	CGCCCGTGCC	GGGACACCGT
	62281	TCGCCCAGCG	GCTCGCCGAG	CTGCCCGACG	CCGACCGCGG	CGCGGCGCTG	ACCCAGCTCG
30	62341	TCTCGGACGC	CACGGCCGCC	GTGCTCGGCC	ACGCCGACGC	CTCCGAGATC	GCGCCGACCA
	62401	CGACGTTCAA	GGACCTCGGC	ATCGACTCGC	TCACCGCGAT	CGAGCTGCGC	AACCGGCTCG
	62461	CGGAGGCGAC	CGGGCTGCGG	CTGAGTGCCA	CGCTGGTGTT	CGACCACCCG	ACACCTCGGG
	62521	TCCTCGCCGC	CAAGCTCCGC	ACCGATCTGT	TCGGCACGGC	CGTGCCACAG	CCCGCGCGGA
	62581	CGGCACGGAC	CCACCACGAC	GAGCCACTCG	CGATCGTCGG	CATGGCGTGC	CGACTGCCCC
35	62641	GCGGGTCCGC	CTCGCCGAG	GACCTGTGGC	AGCTCGTGCG	GTCGCGACCC	GACGCGATCA
	62701	CCGAGTTCCC	CACCGACCGC	GGCTGGGACA	TCGACCGGCT	GTTTCGACCCG	GACCCGGACG
	62761	CCCCCGGCAA	GACCTACGTC	CGGCACGGCG	GCTTCCTCGC	CGAGGCCGCC	GGCTTCGATG
	62821	CCGCGTTCTT	CGGCATCAGC	CCGCGCGAGG	CACGGGCCAT	GGACCCGCGC	CAGCGCGTCA
	62881	TCCTCGAAAC	CTCCTGGGAG	GCGTTCGAGA	ACGCGGGCAT	CGTGCCGGAC	ACGCTGCGCG
40	62941	GCAGCGACAC	CGGCGTGTTT	ATGGGCGCGT	TCTCCCATGG	GTACGGCGCC	GGCGTCGACC
	63001	TGGGCGGGTT	CGGCGCCACC	GCCACGCAGA	ACAGCGTGCT	CTCCGGCCGG	TTGTCTGACT
	63061	TCTTCGGCAT	GGAGGGCCCG	GCCGTACCCG	TGCACACCGC	CTGCTCGTCG	TCGCTGGTCG
	63121	CCCTGCACCA	GGCGGCACAG	GCGCTGCGGA	CTGGAGAATG	CTCGCTGGCG	CTCGCCGGCG
	63181	GTGTACGGT	GATGCCCCACC	CCGCTGGGCT	ACGTGAGTTC	CTGCCGCCAG	CGGGGACTCG
45	63241	CCCCCGACGG	CCGTTGCCAG	GCTTCGCGGG	AAGGCGCCGA	CGGCACGAGC	TTCTCGGAGG
	63301	GCGCCGGCGT	TCTTGTCGTC	GAGCGGCTCT	CCGACGCCGA	GCGCAACGGA	CACACCGTCC
	63361	TCGCGGTCGT	CCGCTCCTCC	GCCGTCAACC	AGGACGGCGC	CTCCAACGGC	ATCTCCGCAC
	63421	CCAACGGCCC	CTCCCAGCAG	CGCGTCATCC	GCCAGGCCCT	CGACAAGGCC	GGGCTCGCCC
	63481	CCGCCGACGT	GGACGTGGTG	GAGGCCACAG	GCACCGGAAC	CCCGCTGGGC	GACCCGATCG
50	63541	AGGCACAGGC	CATCATCGCG	ACCTACGGCC	AGGACCGCGA	CACACCGCTC	TACCTCGGTT
	63601	CGGTCAAGTC	GAACATCGGA	CACACCCAGA	CCACCGCCGG	TGTCGCCCGC	GTCATCAAGA
	63661	TGGTCATGGC	GATGCGCCAC	GGCATCGCGC	CGAAGACACT	GCACGTGGAC	GAGCCGTCTG
	63721	CGCATGTGGA	CTGGACCGAG	GGTGCGGTGG	AACTGCTCAC	CGAGGCGAGG	CCGTGGCCCC
	63781	ACGCGGGACG	CCCGCGCCGC	GCGGCGGTGT	CGTCGCTCGG	TATCAGCGGT	ACGAACGCCC
55	63841	ACGTGATCCT	TGAGGGTGTT	CCCGGGCCGT	CGCGTGTGGA	GCCGTGTGTT	GCCGCTTGG
	63901	TGCCGTTGCC	GGTGTCGGCT	CGGAGTGAGG	CGAGTCTGCG	GGGGCAGGTG	GAGCGGCTGG
	63961	AGGGGTATCT	GCGCGGGAGT	GTGGATGTGG	CCGCGGTGCG	GCAGGGGTTG	GTGCGTGAGC
	64021	GTGCTGTCTT	CGGTACCCGT	GCGGTACTGC	TGGGTGATGC	CCGGGTGATG	GGTGTGGCGG
	64081	TGGATCAGCC	CCGTACGGTG	TTGCTCTTTC	CCGGGCAGGG	TGCTCAGTGG	GTGGGCATGG
60	64141	GTGTGGAGTT	GATGGACCGT	TCTGCGGTGT	TCGCGGCTCG	TATGGAGGAG	TGTGCGCGGG
	64201	CGTTGTTGCC	GCACACGGGC	TGGGATGTGC	GGGAGATGTT	GGCGCGGCCG	GATGTGGCGG
	64261	AGCGGGTGA	GGTGGTCCAG	CCGGCCAGCT	GGGCGGTGCG	GCTCAGCCTG	GCCGCACTGT
	64321	GGCAGGCCCA	CGGGGTCTGA	CCCAGCGCGG	TGATCGGACA	CTCCCAGGGC	GAGATCGCGG
	64381	CGGCGTGCGT	GGCCGGGGCC	CTCAGCCTTG	AGGACGCCCG	CCGCGTGGTG	GCCTTGCGCA

	64441	GCCAGGTCAT	CGCGGCGCGA	CTGGCCGGGG	GGGGAGCGAT	GGCTTCGGTG	GCATTGCCGG
	64501	CCGGTGAGGT	CGGTCTGGTC	GAGGGCGTGT	GGATCGCGGC	CGGTAACGGC	CCCGCCTCGA
	64561	CAGTCGTGGC	CGGCGAGCCG	TCGGCGGTGG	AGGACGTGGT	GACGCGGTAT	GAGACCGAAG
	64621	GCGTGCGAGT	GCGTCGTATC	GCCGTGCACT	ACGCCTCCCA	CACGCCCCAC	GTGGAAACCA
5	64681	TCGAGGACGA	ACTCGCTGAG	GTACTGAAGG	GAGTTGCAGG	GAAGGCCGCG	TCGGTGGCGT
	64741	GGTGGTCGAC	CGTGGACAGC	GCCTGGGTGA	CCGAGCCGGT	GGATGAGAGT	TACTGGTACC
	64801	GGAACCTGCG	TCGCCCCGTC	GCGCTGGACG	CGGCGGTGGC	GGAGCTGGAC	GGGTCCGTGT
	64861	TCGTGGAGTG	CAGCGCCCAT	CCGGTGCTGC	TGCCGGCGAT	GGAACAGGCC	CACACGGTGG
	64921	CGTCGTTCGG	CACCGGTGAC	GGCGGTGGG	AGCGATGGCT	GACGGCGTTG	GCGCAGGCGT
10	64981	GGACCCGTTG	CGCGGCAGTG	GACTGGGACA	CGGTGGTCCA	ACCGGTGCCA	GGCGGCTGTC
	65041	TCGATCTGCC	CACCTACGCG	TTCGAGCGCC	GGCGCTACTG	GCTGGAAGCG	GCCGGTGCCA
	65101	CCGACCTGTC	CGCGGCCGGG	CTGACAGGGG	CAGCACATCC	CATGCTGGCC	GCCATCACGG
	65161	CACACCCGCG	CGACGACGGT	GGTGTGTTC	TCACCGGCCG	GATCTCGTTG	CGCACGCATC
	65221	CTGGCTGGC	TGATCACGCG	GTGCGGGGCA	CGGTCTTGCT	GCCGGGCACG	GCCTTTGTGG
15	65281	AGCTGGTCAT	CCGGGCCGGT	GACGAGACCG	GTTGCGGGAT	AGTGGATGAA	CTGGTCATCG
	65341	AATCCCCCCT	CGTGGTGCCG	GCGACCGCAG	CCGTGGATCT	GTGCGTGACC	GTGGAAGGAG
	65401	CTGACGAGGC	CGGACGGCGG	CGAGTGACCG	TCCACGCCCG	CACCGAAGGC	ACCGGCAGCT
	65461	GGACCCGGCA	CGCCAGCGGC	ACCCTGACCC	CCGACACCCC	CGACACCCCC	AACGCTTCCG
	65521	GTGTTGTCCG	TGCGGAGCCG	TTCTCGCAGT	GGCCACCTGC	CACTGCCGCG	GCCGTCGACA
20	65581	CCTCGGAGTT	CTACTTGCGC	CTGGACGCGC	TGGGCTACCG	GTTCCGGACCC	ATGTTCCGCG
	65641	GAATGCGGGC	TGCTGGCGT	GATGGTGACA	CCGTGTACGC	CGAGGTCCGCG	CTCCCCGAGG
	65701	ACCGTGCCCG	CGACGCGGAC	GGTTTCGGCA	TGCACCCGGC	GCTGCTCGAC	GCGGCTTTCG
	65761	AGAGCGGCAG	CCTGCTCATG	CTGGAATCCG	ACGCGGAGCA	GAGCGTGCAA	CTGCCCTTCT
	65821	CCTGGCACGG	CGTCCGGTTC	CACGCGACGG	GCGCGACCAT	GCTGCGGGTG	GCGGTGCTAC
25	65881	CGGGCCCGGA	CGGCCTCCGG	CTGCATGCCG	CGGACAGCGG	GAACCGTCCC	GTCGCGACGA
	65941	TCGACGCGCT	CGTGACCCGG	TCCCCGGAAG	CGGACCTCGC	GCCCGCCGAT	CCGATGCTGC
	66001	GGGTCGGGTG	GGCCCCGGTG	CCCGTACCTG	CCGGGGCCGG	TCCGTCCGAC	GCGGACGTGC
	66061	TGACGCTGCG	CGGCGACGAC	GCCGACCCCG	TCGGGGAGAG	CCGGGACCTG	ACCACCCGTG
	66121	TTCTCGACGC	GCTGCTCCGG	GCCGACCCGG	CGGTGATCTT	CCAGGTGACC	GGTGGCCTCG
30	66181	CCGCCAAGGC	GGCCGCAAGC	CTGGTCCGCA	CCGCTCAGAA	CGGACAGCCC	GGCCGCTTCT
	66241	TCCTCGTCGA	AACGGACCCG	GGAGAGGTCC	TGGACGGCGC	GAAGCGCGAC	GCGATCGCGG
	66301	CACTCGGCGA	GCCCCATGTG	CGGCTGCGCG	ACGGCCTCTT	CGAGGCAGCC	CGGCTGATGC
	66361	GGGCCACGCC	GTCCCTGACG	CTCCCGGACA	CCGGGTCTGT	GCAGCTGCGG	CCGTCCGCCA
	66421	CCGGTTCCCT	CGACGACCTT	GCCGTCTGTC	CCACCGACGC	CCCGGACCGG	CCGCTCGCGG
35	66481	CCGGCGAGGT	GCGGATCGCG	GTACGCGCGG	GGGGCCTGAA	CTTCCGGGAT	GTCACGGTCG
	66541	CGCTCGGTGT	GGTCGCCGAT	GCGCGTCCCG	TCGGCAGCGA	GGCCCGGGT	GTCGTCTGG
	66601	AGACCGGCCC	CGGTGTGCAC	GACCTGGCGC	CCGGCGACCG	GGTCTGGGG	ATGCTCGCGG
	66661	GCGCCTTCGG	ACCGGTTCGCG	ATCACCGACC	GGCGGCTGCT	CGGCCGGATG	CCGGACGGCT
	66721	GGACGTTCCT	GCAGGCGGCG	TCCGTGATGA	CCGCGTTCGC	GACCGCGTGG	TACGGCCTGG
40	66781	TCGACCTGGC	CGGGCTGCGC	CCCGGCGAGA	AGGTCTTGAT	CCACGCGGCG	GCGACCGGTG
	66841	TCGGCGCGGC	GGCCGTCCAG	ATCGCGCGGC	ATCTGGGCGC	GGAGGTGTAC	GCGACACCA
	66901	GCGCCGCGAA	GCGCCATCTG	GTGGACCTGG	ACGGAGCGCA	TCTGCCGAT	TCCCGCAGUA
	66961	CCGCGTTCGC	CGACGCGTTC	CCGCCGCTCG	ATGTCGTGCT	CAACTCGCTC	ACCGGTGAAT
	67021	TCCTCGACGC	GTCCGTTCGGC	CTGCTCGCGG	CGGGTGGCCG	GTTTCATCGAG	ATGGGGAAGA
45	67081	CGGACATCCG	GCACGCCGTC	CAGCAGCCGT	TCGACCTGAT	GGACGCCGCG	CCCGACCGGA
	67141	TGCAGCGGAT	CATCGTCGAG	CTGCTCGGCC	TGTTGCGCGC	CGACGTGCTG	CACCCGCTGC
	67201	CGGTCCACGC	CTGGGACGTG	CGGCAGGCGC	GGGAGGCGTT	CGGCTGGATG	AGCAGCGGGC
	67261	GTCACACCGG	CAAGCTGGTG	CTGACGGTCC	CGCGGCCGCT	GGATCCCGAG	GGGGCCGTCG
	67321	TCATCACCGG	CGGCTCCGGC	ACCCTCGCCG	GCATCCTCGC	CCGCCACCTG	GGCCACCCCC
50	67381	ACACCTACCT	GCTCTCCCGC	ACCCACCCCC	CCGACACCAC	CCCCGGCACC	CACCTCCCT
	67441	GCGACGTCGG	CGACCCCCAC	CAACTCGCCA	CCACCTCGC	CCGCATCCCC	CAACCCCTCA
	67501	CCGCCGTCTT	CCACACCGCC	GGAACCCCTG	ACGACGCCCT	GCTCGACAAC	CTCACCCCGG
	67561	ACCGCGTCGA	CACCGTCCCT	AAACCCAAAG	CCGACGCCGC	CTGGCACCTG	CACCGGCTCA
	67621	CCCGCGACAC	CGACCTCGCC	GCGTTCGTGC	TCTACTCCGC	GGTCGCCCGC	CTCATGGGCA
55	67681	GCCCGGGGCA	GGGCAACTAC	GTCCGCGGCA	ACGCGTTCCT	CGACGCGCTC	GCCGAACACC
	67741	GCCGTGCGCA	AGGGCTGCCC	GCGCATGCCC	TGCGATGGGG	CATGTGGGCG	GACGTACGCG
	67801	CGCTCACCGC	GAAACTCACC	GACGCGGACC	GCCAGCGCAT	CCGGCGCAGC	GGATTCCCGC
	67861	CGTTGAGCGC	CGCGGACGGC	ATGCGGCTGT	TCGACGCGGC	GACGCGTACC	CCGGAACCGG
	67921	TCGTCTGCGC	GACGACCGTC	GACCTCACCC	AGCTCGACGG	CGCCGTGCGC	CCGTTGCTCC
60	67981	GCGGTCTGGC	CGCGCACCGG	GCCGGGCCGG	CGCGCACGGT	CGCCCCGAAC	GCCGGCGAAG
	68041	AGCCCCGTCG	CGTGCGTCTT	GCCGGGCGTA	CCGCCGCCGA	GCAGCGGCGC	ATCATGACGG
	68101	AGGTCGTGCT	CCGCCACGCG	GCCGCGGTCC	TCGCGTACGG	GCTGGGCGAC	GCGGTGCGCG
	68161	CGGACGCTCC	GTTCCGCGAG	CTCGGTTTCG	ATTCGCTGAC	CGCGGTGCGC	CTGCGCAATC
	68221	GGCTCGCGGC	CGAGACGGGG	CTGCGGCTGC	CGACGACGCT	GGTGTTCAGC	CACCCGACGG

68281 CGGAGGCGCT CACCGCCAC CTGCTCGACC TGATCGACGC TCCACCGCC CGGATCGCCG
 68341 GGGAGTCCCT GCGCGCGTG ACGCCGCTC CCGTGGCGGC CGCGCGGGAC CAGGACGAGC
 68401 CGATCGCCAT CGTGGCGATG GCGTGCCGGC TGCCCGGTGG TGTGACGTCG CCCGAGGACC
 5 68461 TGTGGCGGCT CGTCGAGTCC GGCACCGACG CGATCACCAC GCCTCCTGAC GACCGCGGCT
 68521 GGGACGTCGA CGCGCTGTAC GACGCGGACC CGGACGCGGC CGGCAAGGCG TACAACCTGC
 68581 GGGGCGG'FTA CCTGGCCGGG GCGGCGGAGT TCGACGCGGC GTTCTTCGAC ATCAGTCCGC
 68641 GCGAAGCGCT CGGCATGGAC CCGCAGCAAC GCCTGCTGCT GGAACCGGC TGGGAGGCGA
 68701 TCGAGCGCGG CCGGATCAGT CCGGCGTCGC TCCGCGGCCG GGAGGTCGGC GTCTATGTCG
 10 68761 GTGCGGCCGC GCAGGGCTAC GGGCTGGGCG CCGAGGACAC CGAGGGCCAC GCGATCACC
 68821 GTGGTTCCAC GAGCCTGCTG TCCGGACGGC TGGCGTACGT GCTCGGGCTG GAGGGCCCGG
 68881 CCGTACCCGT GGACACGGCG TGCTCGTCGT CTCTGGTCGC GCTGCATCTG GCGTGCCAGG
 68941 GGTGCGCCT GGGCGAGTGC GAATCGCTC TGGCCGGAGG GGTCTCCGTA CTGAGTTCGC
 69001 CGGCCGCGTT TCCCGCCGAGT CCGGGCTCGC GCGGGCTCGC GGCGGAGG ATCTGCAAGT
 69061 CGTTCGGCGC GGGCGCGGAC GGCACGACGT GGTCCGAGGG CGTGGGCGTG CTCGTACTGG
 15 69121 AACGGCTCTC CGACGCCGAG CCGCTCGGGC ACACCGTGCT CGCCGTCGTC CGCGGCAGCG
 69181 CCGTACCGTC CGACGGCGCC TCCAACGGCC TCACCGCGCC GAACGGGCTC TCGCAGCAGC
 69241 GGGTCATCCG GAAGGCGCTC GCGCGGCCG GGTGACCGG CGCCGACGTG GACGTCGTCG
 69301 AGGGGACCGG CACCGGCACC CCGCTCGGC ACCCGGTCGA GGCGGACGCG CTGCTCGCGA
 20 69361 CGTACGGGCA GGACCGTCCG GCACCGGCTC GGTGGGCTC GCTGAAGTCG AACATCGGAC
 69421 ATGCCACGGC CGCGGCCGGT GTCGCGGGCG TCATCAAGAT GGTGCAAGCG ATCGGCGCGG
 69481 GCACGATGCC GCGGACGCTG CATGTGGAGG AGCCCTCGCC CGCCGTCGAC TGGAGCACC
 69541 GACAGGTGTC CCTGCTCGGC TCCAACGGC CCTGGCCGGA CGACGAGCGT CCGCGCCGGG
 69601 CGGCCGTCTC CGCGTTCGGG CTCAGCGGGA CGAACGCGCA CGTCATCCTG GAACAGCACU
 69661 GTCCGGCGCC CGTGGCGTCC CAGCCGCCCC GCGCGCCCCG TGAGGAGTCC CAGCCGCTGC
 25 69721 CGTGGGTGCT CTCCGCGCGG ACTCCGCGCG CGTGGCGGGC CCAGGCGGGC CGGCTGCGCG
 69781 ACCACCTCGC GCGGCGACCG GACGCGGATC CGTTGGACAT CCGGTACGCG CTGGCCACCA
 69841 GCGCGCCCA GTTCGCCCCAC CGTGCCGCGG TCGTCGCCAC CACCCCGGAC GGATTCGCTG
 69901 CCGCGCTCGA CGGCCTCGCG GACGGCGCGG AGGCGCCCG AGTCGTCACC GGGACCGCTC
 69961 AGGAGCGGGC CGTCGCCTTC CTCTTCGACG GCCAGGGCGC CCAGCGCGCC GGAATGGGGC
 30 70021 GCGAGCTCCA CCGCCGGTTC CCGCTCTTCG CCGCCGCGTG GGACGAGGTC TCCGACGCGT
 70081 TCGGCAAGCA CCTCAAGCAC TCCCCACGG ACGTCTACCA CGGCGAACAC GCGCTCTCG
 70141 CCCATGACAC CCTGTACGTC CAGGCGGGC TGTTACGCT CGAAGTGGCG CTGCTGCGGC
 70201 TGTGGAGCA CTGGGGGGTG CCGCCGACG TGCTCGTCGG GCATCCGTC GCACTCCGTA
 70261 CCGCGGCGTA CGCGCGGGG GTGCTACCC TGGCGGACGC GACGGAGTTG ATCGTGGCCC
 35 70321 GGGGGCGGGC GCTGCGGGCG CTGCGCCCCG GGGCGATGCT CGCCGTCGAC GGAAGCCCGG
 70381 CGGAGGTCGG CCGCCGACG GATCTGGACA TCGCCGCGGT CAACGGCCCG TCCGCCGTGG
 70441 TGCTCGCCGG TTCGCGGAC GATGTGGCGG CGTTCGAACG GGAGTGGTCG GCGGCCGGG
 70501 GCGGCACGAA ACGGCTCGAC GTCGGGCAGC CGTTCCACTC CCGGCACGTC GACGGTGGCG
 70561 TCGACGGCTT CCGTACGGTG CTGGAGTCGC TCGGTTTCGG CGCGGCGCG CTGCCGTGG
 40 70621 TGTCCACGAC GACGGGCGCG GACGCGCGG ACGACCTCAT AACGCCCGC CACTGGCTGC
 70681 GCCATGCGCG TCGGCGGGTG CTGTTCTCGG ATGCGTCCG GGAGCTGGCC GACCGCGGCG
 70741 TCACCACGTT CGTGGCCGTC GGCCCTCCG GCTCCCTGGC GTCGGCGCG GCGGAGAGCG
 70801 CCGGGGAGGA CGCCGGGACC TACCACGCGG TGCTGCGCGC CCGGACCGGT GAGGAGACCG
 70861 CCGCGCTGAC CGCCCTCGCC GAGCTGCACG CCCACGGCGT CCGGTCGAC CTGGCCGCGG
 45 70921 TACTGGCCGG TGGCCGGCCA GTGGACCTTC CCGTGTACGC GTTCCAGCAC GGTTCCTACT
 70981 GGCCTGGCCC GCGCGTGGCG GGGGCGCCG CCACCGTGGC GGACCGCGG GGTCCGCGG
 71041 AGTCCGAGCC GGAGGACCTC ACCGTCGCGG AGATCGTCCG TCGGCGCACC GCGGCGCTGC
 71101 TCGGCGTCAC GGACCCCGCC GACGTCGATG CGGAAGCGAC GTTCTTCGCG CTCGGTTTCG
 71161 ACTCACTGGC GGTGCAGCGG CTGCGCAACC AGCTCGCCTC GGCAACCGGG CTGGACCTGC
 50 71221 CCGCGGCCGT CCTGTTTCGAC CACGACACCC CCGCCGCGCT CACCGCGTTC CTCCAGGACC
 71281 GGATCGAGGC CGGCCAGGAC CCGATCGAGG CCGGCGAGGA CGACGACGCG CCCACCGTGC
 71341 TCTCGCTCCT GGAGGAGATG GAGTCGCTCG ACGCCGCGGA CATCGCGCG CACCCGGCCC
 71401 CCGAGCGTGC GGCCATCGCC GATCTGCTCG ACAAGCTCGC CCATACCTGG AAGGACTACC
 71461 GATGAGCACC GATACGCACG AGGGAACGCC GCGCGCCGGC CGCTGCCCAT TCGCGATCCA
 55 71521 GGACGGTCAC CGCGCCATCC TGGAGAGCGG CACGGTGGGT TCGTTCGACC TGTTCCGGCGT
 71581 CAAGCACTGG CTGGTCGCCG CCGCCGAGGA CGTCAAGCTG GTCACCAACG ATCCGCGGTT
 71641 CAGCTCGGCC GCGCCGTCCG AGATGCTGCC CGACCGGCGG CCGGCTGGT TCTCCGGGAT
 71701 GGAATCACC GAGCACAACC GCTACCGGCA GAAGATCGCG GGGGACTTCA CACTGCGCGC
 71761 GGGCGCAAG CCGGAGGACT TCGTCGCCGA GCGCGCCGAC GCCTGCCTGC ACGACATCGA
 60 71821 GCGCGCGGGA CCGGCAACC ACCTCATCCC CGGGTACGCC AAGCGGCTGC CTTCCCTCGT
 71881 CATCAACGCG CTGTACGGG TCACCCCTGA GGAGGGGGC GTGCTGGAGG CACGGATGCG
 71941 CGACATCACC GGCTCGGCC ATCTGGACAG CGTCAAGACG CTGACCGACG ACTTCTTCGG
 72001 GCACGCGCTG CCGCTGGTCC GCGCGAAGCG TGACGAGCGG GGCGAGGACC TGCTGCACCG
 72061 GCTGGCCTCG GCGGACGACG GCGAGATCTC GCTCAGCGAC GACGAGGCGA CCGGCGTGTT

	72121	CGCGACGCTG	CTGTTCGCCG	GCCACGACTC	GGTGCAGCAG	ATGGTCGGCT	ACTGCCTCTA
	72181	CGCACTGCTC	AGCCACCCCG	AGCAGCAGGC	GGCGCTGCGC	GCGCGCCCGG	AGCTGGTCTGA
	72241	CAACGCGGTC	GAGGAGATGC	TCCGTTTCCT	GCCCGTCAAC	CAGATGGGCG	TACCGCGCGT
	72301	CTGTGTCGAG	GACGTTCGATG	TGCGGGGCGT	GCGCATCCGT	GCGGGCGACA	ACGTGATCCC
5	72361	GCTCTACTCG	ACGGCCAACC	GCGACCCCGA	GGTGTTCCTG	CAGCCCGACA	CCTTCGATGT
	72421	GACGCGCCCG	CTGGAGGGCA	ACTTCGCGTT	CGGCCACGGC	ATTACAAGT	GTCCCGGCCA
	72481	GCACATCGCC	CGGGTGCTCA	TCAAGGTCGC	CTGCCTGCGG	TTGTTTCGAGC	GTTTCCCGGA
	72541	CGTCCGGGCTG	GCCGGCGACG	TGCCGATGAA	CGAGGGGCTC	GGGCTGTTCA	GCCCGGCCGA
10	72601	GCTGCGGGTC	ACCTGGGGGG	CGGCATGAGT	CACCCGGTGG	AGACGTTGCG	GTTGCCGAAC
	72661	GGGACGACGG	TCGCGCACAT	CAACGCGGGC	GAGGCGCAGT	TCCTCTACCG	GGAGATCTTC
	72721	ACCCAGCGCT	GCTACCTGCG	CCACGGTGTC	GACCTGCGCC	CGGGGGACGT	GGTGTTCGAC
	72781	GTCGGCGCGA	ACATCGGCAT	GTTACGCTT	TTCGCGCATC	TGGAGTGTCC	TGGTGTGACC
	72841	GTGCACGCCT	TCGAGCCCCG	GCCCGTGCCG	TTCGCGGCGC	TGCGGGCGAA	CGTGACGCGG
	72901	CACGGCATCC	CGGGCCAGGC	GGACAGTGTC	GCGGTCTCCG	ACAGCTCCGG	CACCCGGAAG
15	72961	ATGACCTTCT	ATCCCGACGC	CACGCTGATG	TCCGGTTTCC	ACGCGGATGC	CGCGGCCCGG
	73021	ACGGAGCTGT	TGCGCACGCT	CGGCCTCAAC	GGCGGCTACA	CCGCGGAGGA	CGTCGACACC
	73081	ATGCTCGCGC	AACTGCCCCA	CGTCAGCGAG	GAGATCGAAA	CCCCTGTGGT	CCGGCTCTCC
	73141	GACGTTCATCG	CGGAGCGCGG	TATCGAGGCC	ATCGGCCTGC	TGAAGGTCGA	CGTGGAGAAG
	73201	AGCGAACGGC	AGGTCTTCGC	CGGCCTCGAG	GACACCGACT	GGCCCCGTAT	CCGCCAGGTC
20	73261	GTCGCGGAGG	TCCACGACAT	CGACGGCGCG	CTCGAGGAGG	TCGTACACGT	GCTCCGCGGC
	73321	CATGGCTTCA	CCGTGGTTCG	CGAGCAGGAA	CCGCTGTTCG	CCGGCACGGG	CATCCACCAG
	73381	GTCGCCGCGC	GGCGGGTGGC	CGGCTGAGCG	CCGTCGGGGC	CGCGCCCGTC	CGCACCGGCG
	73441	GCCGCGGTGC	GGACGGCGGC	TGACGCGCGC	TCGGACAGTT	CCTTGGGCGA	TTGTGACGG
	73501	CCCTTCACCC	CCAGCTTGCG	GAACACGTTG	GTGAGGTGCT	GTTCCACCGT	GTTGGAGGTC
25	73561	ACGAACAGCT	GGCTGGCGAT	CTCCTTGTTG	GTGCGCCCGA	CCGCGGCGTG	CGACGCCACC
	73621	CGCCGCTCCG	CCTCGGTCAG	CGATGTGATC	CGCTGCGCCG	GCGTCACGTC	CTGGGTGCCG
	73681	TCCGCGTCCG	AGGACTCCCC	ACCGAGCCGC	CGGAGGAGCG	GCACGGCTCC	GCACTGGGTC
	73741	GCGAGGTGCC	GTGCGCGGCG	GAACAGTCCC	CGCGCACGGC	TGTGCCGCGG	GAGCATGCCG
	73801	CACGCTTCGC	CCATGTTCGC	GAGGACGCGG	CCCAGCTCGT	ACTGGTCGCG	GCACTGATG
30	73861	AGCAGATCGG	CGGCCTCGTC	GAGCAGTTTC	ATCCGCTTGG	CCGCGGCACT	GTAGGCGGCC
	73921	TGCACCCGCA	GCGTCATCAC	CCGCGCCCGG	GACCCCATCG	GCCGGGACAG	CTGCTCGGAG
	73981	ATGAGCCTCA	GCCCCCTCGT	ACGGCCGCGG	CCGAGCAGCA	GAAGCGCTTC	GGCGGCGTCG
	74041	ACCCGCCACA	GGGCCAGGCC	CGGCACGTCG	ACGGACCAGC	GTGCGATCCG	CTCCCCGCAG
	74101	TCCCGGAACG	CGTTGTACGC	CGCCCGGTAC	CGCCCGGCCG	CGAGATGGTG	TTGCCACCGG
35	74161	GCCCAGACCA	TGTGCAGTCC	GAAGAGGCTG	TCGGAGGTCT	CCTCCGGCAA	CGGCTCGGCG
	74221	AGCCACCGCT	CCGCCCGGTC	CAGGTGCGCC	AGTCGGATCG	CGGCGGCCAC	GGTGTGCTC
	74281	AGCGGCAATG	CGGCGGCCAT	CCCCCAGGAG	GGCACGACCC	GGGGGGCGAG	CGCGGCCTCG
	74341	CCGCATTCTGA	CGGCGGCGGT	CAGGTGCGCG	CGGCGCAGCG	CGGCCTCGGC	GCGGAACCCC
	74401	GCGTGGACCG	CCTCGTCCGC	CGGGGTCCGC	ATGTTGTCTG	CACCGGCCAG	CTTGTGACCC
40	74461	CAGGACTGGA	CGGCATCGGT	GTCTCTCGGC	TAGAGCAGGG	CCAGCAACGC	CATCATGGTC
	74521	GTGGTCCGGT	CCGTCTGTGAC	CCGGGAGTGC	TGGAGCACGT	ACTCGGCTTT	GGCCTCGGCC
	74581	TGTTCCGACC	AGCCGCGCAG	CGCCTTGCTC	AGGGCCTTGT	CGGCGACGGC	CGGCTCGCCG
	74641	ACGGTCCCGG	AAAACGAGGC	GACCTCGTCC	TCGGCCGCGG	GATCGGCGCG	ACGCGGCGGA
	74701	TCGGCCGCGC	CGGGATAGAT	CAGCGCGAGG	GACAGGTCCG	CGACGCGCAG	GTGCGCCCGG
45	74761	CCCTGCTCGC	TCGGGGCGGC	GGAGCGCTGG	GCCGCCAGGA	CCTCGGCGGC	CTCGCCCGGC
	74821	CGCCCGTCCA	TCGCCAGCCA	GCAGGCGAGC	GACACGGCGT	GCTCGCTGGA	GAGGAGCCGT
	74881	TCCCGCGACG	CGGTGAGCAG	CTCGGGCACA	TGCCGGCCCG	ATCTGGCGGG	ATCGCAGAGC
	74941	CGCTCGATGG	CGGCGGTGTC	GACGCGCAGT	GCGGCGTGGA	CGGCGGGGTC	GTGCGAGGCC
50	75001	CGGTAGGCGA	ACTCCAGGTA	GGTGACGGCC	TCGTGAGGCT	CGCCGCGCAG	GTAGGTGCTC
	75061	CGCGCGGCGT	CGGTGAACAG	CCCGGCGACC	TCGGCGCCGT	GCACCCGGCC	GGTACCCATC
	75121	TGGTGGCGGG	CGAGCACCTT	GCTGGCCACG	CCGCGGTCCC	GCAGCAGTTC	CAGCGCCAGC
	75181	TCGTGACGGC	CACGCGGCTC	GGCGGCGGAG	AGGTGCTCGA	GTACGACGGA	GCGGGCCGCG
	75241	GGGTGCGGGA	ACCGCCCTTC	CCGCAGCAGC	CGCCCCCTCG	CCAGCTGTTT	GTGGGCCTGC
	75301	TCGACCGCCT	CGGTGTTCGAG	GCCGGTCATC	CGCTGGACGA	GGGTGAGTTC	GACACTCTCG
55	75361	CCGAGCACGG	CGGAAGCTCG	GGCGACGCTC	AGCGCGGCGG	GGCCGCAACG	ATAGAGCGAC
	75421	CCGAGGTAGG	CGAGCCGGTA	GCCCCGCCCC	GCGACCACTT	CCAGGCAACC	TAGGATCCGT
	75481	GTCCGTGCCT	CCCGGATGTC	GTCGATCAGG	CCGTGGCCGA	GGAGCAGGTT	GCCGCGGGTC
	75541	GCCCGGAACG	CCTGGGGCAC	CACGTCTGTC	TGCGCGTCCT	GGCCGAGGTG	CCGGCGCACG
	75601	AGTTCCGGTG	TCTGCGCCTC	GGTGAGCGGG	CGCAGCGCGA	TCTCTGGTGA	GTGGCGCAGA
60	75661	CTCAGCAGTG	CCGCCCCGAA	TTGGGAGTGG	GCGGGCGTCG	GCCGGAGCAG	CTCGGTGAGC
	75721	ACGATGGCGA	CACGGGCCCC	GCTGATGCGG	CGCGCGAGGT	GGAGCAGGCA	GCGCAGCGAC
	75781	GGCGCGTCGG	CGTGGTGCAC	GTCGTCGATG	CCGATCAGTA	CGGGCCGCTC	CGCGCGGAGC
	75841	GTCAGCACCG	TGCGGGTGAG	TTCGTTCCCC	AGGCGGTTGT	CGACGTCGGC	CGCGAGGTTT
	75901	TCGCAACGATG	CCGTGACCGG	GACCAGCTCC	GGTGTCCGGG	CGGCCAGCTC	GGGCTGGTCC

75961 AGGAGCTGGC CGAGCATGCC GTACGGCAGG GCCCGCTCCT CCATGGAGCA CACCCGCGCA
 76021 AGGGTGACGA AGCCGGCCTT GGCCGCGGCG GCGTCGAGGA GTTCGGTCTT GCCGCAGGCG
 76081 ATCGGCCCCG TGACGGCGGC GACGACGCC CGCCCGCCCC CCGCTCGGGT GAGCGCCCGG
 76141 TGGAGGGAAC CGAACTCGTC ATCGCGGGCG ATCAGGTCTG GGGGAGATAA GCGCGCTATC
 5 76201 ACGAATGGAA CTACCTCGCG ACCGTCGTGG AAACCCATAG GCATCACATG GCTTGTGAT
 76261 CTGTACGGCT GTGATTCAGC CTGGCGGGAT GCTGTGCTAC AGATGGGAAG ATGTGA²CTA
 76321 GGGCCGTGCC GTTCCCTCAG GAGCCGACCG CCCCCGGCGC CACCCGCCGT ACCCCCTGGG
 76381 CCACCAGCTC GCGGACCCGC TCCTGGTGGT CGACGAGGTA GAAGTGCCCG CCGGGGAAGA
 76441 CCTCCACCGT GGTCGGCGCG GTCGTGTGCC CGGCCAGGC GTGGGCCTGC TCCACCGTCG
 10 76501 TCTTCGGATC GTCGTCACCG ATGCACACCG TGATCGGCGT CTCCAGCGGC GCGCGGGCT
 76561 CCCACCGGTA CGTCTCCGCC GCGTAGTAGT CCGCCCGCAA CGGCGCCAGG ATCAGCGCGC
 76621 GCATTTTCGTC GTCCGCCATC ACATCGGCGC TCGTCCCGCC GAGGCCGATG ACCGCCGCCA
 76681 GCAGCTCGTC GTCGGACGCG AGGTGGTCTT GGTCGGCGCG CGGCTGCGAC GGCGCCCGCC
 76741 GGCCCGAGAC GATCAGGTGC GCCACCGGGA GCCGCTGGGC CAGCTCGAAC GCGAGTGTG
 15 76801 CGCCCATGCT GTGGCGGAAC AGCACCAGCG GACGGTCCAG CCCCGGCTTC AACGCCTCGG
 76861 CCACGAGGCC GGCGAGAACA CGCAGGTGCG GCACCGCCTC CTCGTCGCGG CGGTCTGGC
 76921 GGCCGGGGTA CTGCACGGCG TACACGTCCG CCACCGGGGC GAGCGCACGG GCCAGCGGAA
 76981 GGTAGAACGT CGCCGATCCG CCGGCGTGGG GCAGCAGCAC CACCCGTACC GGGGCCTCGG
 77041 GCGTGGGGAA GAACTGCCGC AGCCAGAGTT CCGAGCTCAC CGCACCCCT CCGCCGCGAC
 20 77101 CTGGGGAGCC CGGAACCGG TGATCTCGCG CAAGTGCTTC TCCCGCATCT CCGGGTCGGT
 77161 CACGCCCCAT CCTCCTCCG GCGCCAGACA GAGGACCCG ACTTTGCCGT TGTGCACATT
 77221 GCGATGCACA TCGCGGACCG CCGACCCGAC GTCGTCGAGC GGTAGGTCA CCGCAGCGT
 77281 CGGGTGCACC ATCCCTTGC AGATCAGGCG GTTCGCCTCC CACGCCTCAC GATAGTTGCG
 77341 GAAGTGGGTA CCGATGATCC GCTTCACGGA CATCCACAGG TACCGATTGT CAAAGGCGTG
 25 77401 CTCGTATCCC GAGGTTGACG CGCAGGTGAC GATCGTGCCA CCCCACGTG TCACGT³GAC
 77461 ACTCGCGCCG AACGTCGCGC GCCCCGGGTG CTCGAACACG ATGTGGGGAT CGTCACCGCC
 77521 GGTCAGCTCC CGGATC

Those of skill in the art will recognize that, due to the degenerate nature of the
 30 genetic code, a variety of DNA compounds differing in their nucleotide sequences can be
 used to encode a given amino acid sequence of the invention. The native DNA sequence
 encoding the FK-520 PKS of *Streptomyces hygroscopicus* is shown herein merely to
 illustrate a preferred embodiment of the invention, and the present invention includes
 DNA compounds of any sequence that encode the amino acid sequences of the
 35 polypeptides and proteins of the invention. In similar fashion, a polypeptide can typically
 tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid
 sequence without loss or significant loss of a desired activity. The present invention
 includes such polypeptides with alternate amino acid sequences, and the amino acid
 sequences shown merely illustrate preferred embodiments of the invention.

40 The recombinant nucleic acids, proteins, and peptides of the invention are many
 and diverse. To facilitate an understanding of the invention and the diverse compounds
 and methods provided thereby, the following general description of the FK-520 PKS
 genes and modules of the PKS proteins encoded thereby is provided. This general
 description is followed by a more detailed description of the various domains and
 45 modules of the FK-520 PKS contained in and encoded by the compounds of the
 invention. In this description, reference to a heterologous PKS refers to any PKS other
 than the FK-520 PKS. Unless otherwise indicated, reference to a PKS includes reference

to a portion of a PKS. Moreover, reference to a domain, module, or PKS includes reference to the nucleic acids encoding the same and vice-versa, because the methods and reagents of the invention provide or enable one to prepare proteins and the nucleic acids that encode them.

5 The FK-520 PKS is composed of three proteins encoded by three genes designated *fkfA*, *fkfB*, and *fkfC*. The *fkfA* ORF encodes extender modules 7 - 10 of the PKS. The *fkfB* ORF encodes the loading module (the CoA ligase) and extender modules 1 - 4 of the PKS. The *fkfC* ORF encodes extender modules 5 - 6 of the PKS. The *fkfP* ORF encodes the NRPS that attaches the pipecolic acid and cyclizes the FK-520
10 polyketide.

 The loading module of the FK-520 PKS includes a CoA ligase, an ER domain, and an ACP domain. The starter building block or unit for FK-520 is believed to be a dihydroxycyclohexene carboxylic acid, which is derived from shikimate. The recombinant DNA compounds of the invention that encode the loading module of the
15 FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of methods and in a variety of compounds. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for the loading module of the
20 heterologous PKS is replaced by the coding sequence for the FK-520 loading module, provides a novel PKS coding sequence. Examples of heterologous PKS coding sequences include the rapamycin, FK-506, rifamycin, and avermectin PKS coding sequences. In another embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the
25 coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

 In another embodiment, a portion of the loading module coding sequence is utilized in conjunction with a heterologous coding sequence. In this embodiment, the invention provides, for example, either replacing the CoA ligase with a different CoA
30 ligase, deleting the ER, or replacing the ER with a different ER. In addition, or alternatively, the ACP can be replaced by another ACP. In similar fashion, the corresponding domains in another loading or extender module can be replaced by one or more domains of the FK-520 PKS. The resulting heterologous loading module coding sequence can be utilized in conjunction with a coding sequence for a PKS that
35 synthesizes FK-520, an FK-520 derivative, or another polyketide.

The first extender module of the FK-520 PKS includes a KS domain, an AT domain specific for methylmalonyl CoA, a DH domain, a KR domain, and an ACP domain. The recombinant DNA compounds of the invention that encode the first extender module of the FK-520 PKS and the corresponding polypeptides encoded
5 thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 first extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the first extender module of the FK-520 PKS or the latter is
10 merely added to coding sequences for modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the first extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

15 In another embodiment, all or only a portion of the first extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting either the DH or KR or both; replacing the
20 DH or KR or both with another DH or KR; and/or inserting an ER. In replacing or inserting KR, DH, and ER domains, it is often beneficial to replace the existing KR, DH, and ER domains with the complete set of domains desired from another module. Thus, if one desires to insert an ER domain, one may simply replace the existing KR and DH domains with a KR, DH, and ER set of domains from a module containing such
25 domains. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a gene for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous first extender module coding
30 sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the first extender module of the FK-520 PKS.

In an illustrative embodiment of this aspect of the invention, the invention
35 provides recombinant PKSs and recombinant DNA compounds and vectors that encode

such PKSs in which the KS domain of the first extender module has been inactivated. Such constructs are especially useful when placed in translational reading frame with the remaining modules and domains of an FK-520 or FK-520 derivative PKS. The utility of these constructs is that host cells expressing, or cell free extracts containing, the PKS
5 encoded thereby can be fed or supplied with N-acylcysteamine thioesters of novel precursor molecules to prepare FK-520 derivatives. See U.S. patent application Serial No. 60/117,384, filed 27 Jan. 1999, and PCT patent publication Nos. US97/02358 and US99/03986, each of which is incorporated herein by reference.

The second extender module of the FK-520 PKS includes a KS, an AT specific
10 for methylmalonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the second extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 second extender module is inserted into a DNA compound that comprises the
15 coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the second extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the second
20 extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the second extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid
25 module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of
30 these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous second extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-
35 520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding

domains in a module of a heterologous PKS can be replaced by one or more domains of the second extender module of the FK-520 PKS.

The third extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the third extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 third extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the third extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the third extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the third extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous third extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the third extender module of the FK-520 PKS.

The fourth extender module of the FK-520 PKS includes a KS, an AT that binds ethylmalonyl CoA, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the fourth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In

one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fourth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fourth extender
5 module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the fourth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS
10 that produces an FK-520 derivative.

In another embodiment, a portion of the fourth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the ethylmalonyl CoA specific AT with a malonyl CoA, methylmalonyl CoA, or 2-hydroxymalonyl CoA
15 specific AT; and/or deleting the inactive DH, inserting a KR, a KR and an active DH, or a KR, an active DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, a PKS for a polyketide other than FK-520, or from
20 chemical synthesis. The resulting heterologous fourth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fourth extender module of the FK-520 PKS.

25 As illustrative examples, the present invention provides recombinant genes, vectors, and host cells that result from the conversion of the FK-506 PKS to an FK-520 PKS and vice-versa. In one embodiment, the invention provides a recombinant set of FK-506 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by
30 those for the AT domain of the fourth extender module of the FK-520 PKS. This recombinant PKS can be used to produce FK-520 in recombinant host cells. In another embodiment, the invention provides a recombinant set of FK-520 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of

the fourth extender module of the FK-506 PKS. This recombinant PKS can be used to produce FK-506 in recombinant host cells.

Other examples of hybrid PKS enzymes of the invention include those in which the AT domain of module 4 has been replaced with a malonyl specific AT domain to provide a PKS that produces 21-desethyl-FK520 or with a methylmalonyl specific AT domain to provide a PKS that produces 21-desethyl-21-methyl-FK520. Another hybrid PKS of the invention is prepared by replacing the AT and inactive KR domain of FK-520 extender module 4 with a methylmalonyl specific AT and an active KR domain, such as, for example, from module 2 of the DEBS or oleandolide PKS enzymes, to produce 21-desethyl-21-methyl-22-desoxo-22-hydroxy-FK520. The compounds produced by these hybrid PKS enzymes are neurotrophins.

The fifth extender module of the FK-520 PKS includes a KS, an AT that binds methylmalonyl CoA, a DH, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the fifth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fifth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fifth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS. In another embodiment, a DNA compound comprising a sequence that encodes the fifth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fifth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one or both of the DH and KR; replacing any one or both of the DH and KR with either a KR and/or DH; and/or inserting an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fifth extender module coding sequence

can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fifth extender module of the FK-520 PKS.

5 In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH domain of the fifth extender module have been deleted or mutated to render the DH non-functional. In one such mutated gene, the KR and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. The resulting PKS genes code for the
10 expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-19 to C-20 double bond of FK-520 and has a C-20 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant fifth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment,
15 the present invention provides a recombinant FK-520 PKS that contains both this fifth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that
20 express this recombinant PKS and so synthesize the corresponding (lacking the C-19 to C-20 double bond of FK-506 and having a C-20 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH domain of module 5 has been deleted or otherwise rendered inactive and thus produces this novel polyketide.

25 The sixth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the sixth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes
30 the FK-520 sixth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the sixth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In
35 another embodiment, a DNA compound comprising a sequence that encodes the sixth

extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

5 In another embodiment, a portion of the sixth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, 10 DH, and ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous sixth extender module coding 15 sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the sixth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant 20 FK-520 PKS genes in which the coding sequences for the DH and ER domains of the sixth extender module have been deleted or mutated to render them non-functional. In one such mutated gene, the KR, ER, and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. This can also be accomplished by simply replacing the coding sequences for extender module six with those for an 25 extender module having a methylmalonyl specific AT and only a KR domain from a heterologous PKS gene, such as, for example, the coding sequences for extender module two encoded by the *eryAI* gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that has a C-18 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant 30 activity. This recombinant sixth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this sixth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of 35 the FK-506 PKS. The invention also provides recombinant host cells derived from FK-

506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (having a C-18 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH and ER domains of module 6 have
5 been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The seventh extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the seventh extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of
10 applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 seventh extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the seventh extender module of the FK-520 PKS or the latter is merely added to coding
15 sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the seventh extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

20 In another embodiment, a portion or all of the seventh extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting the KR, the DH, and/or the ER; and/or replacing the
25 KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous
30 seventh extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the seventh extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the seventh extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes
5 code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-15 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant seventh extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an
10 illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this seventh extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-
15 506 but that express this recombinant PKS and so synthesize the corresponding (C-15-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 7 has been replaced and thus produces this novel polyketide.

In another illustrative embodiment, the present invention provides a hybrid PKS
20 in which the AT and KR domains of module 7 of the FK-520 PKS are replaced by a methylmalonyl specific AT domain and an inactive KR domain, such as, for example, the AT and KR domains of extender module 6 of the rapamycin PKS. The resulting hybrid PKS produces 15-desmethoxy-15-methyl-16-oxo-FK-520, a neurotrophin compound.

25 The eighth extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the eighth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520
30 eighth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the eighth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another
35 embodiment, a DNA compound comprising a sequence that encodes the eighth extender

module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the eighth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting or replacing the KR; and/or inserting a DH or a DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous eighth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the eighth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the eighth extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-13 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant eighth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this eighth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-13-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 8 has been replaced and thus produces this novel polyketide.

The ninth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the ninth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 ninth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the ninth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the ninth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the ninth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous ninth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the ninth extender module of the FK-520 PKS.

The tenth extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, and an ACP. The recombinant DNA compounds of the invention that encode the tenth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 tenth extender module is inserted into a DNA compound that comprises the coding sequence

for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the tenth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a
5 DNA compound comprising a sequence that encodes the tenth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the tenth extender module coding
10 sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or inserting a KR, a KR and DH, or a KR, DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP.
15 In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous tenth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that
20 synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the tenth extender module of the FK-520 PKS.

The FK-520 polyketide precursor produced by the action of the tenth extender module of the PKS is then attached to pipecolic acid and cyclized to form FK-520. The
25 enzyme FkbP is the NRPS like enzyme that catalyzes these reactions. FkbP also includes a thioesterase activity that cleaves the nascent FK-520 polyketide from the NRPS. The present invention provides recombinant DNA compounds that encode the *fkbP* gene and so provides recombinant methods for expressing the *fkbP* gene product in recombinant host cells. The recombinant *fkbP* genes of the invention include those in which the
30 coding sequence for the adenylation domain has been mutated or replaced with coding sequences from other NRPS like enzymes so that the resulting recombinant FkbP incorporates a moiety other than pipecolic acid. For the construction of host cells that do not naturally produce pipecolic acid, the present invention provides recombinant DNA compounds that express the enzymes that catalyze at least some of the biosynthesis of
35 pipecolic acid (see Nielsen *et al.*, 1991, *Biochem.* 30: 5789-96). The *fkbL* gene encodes a

homolog of RapL, a lysine cyclodeaminase responsible in part for producing the
pipecolate unit added to the end of the polyketide chain. The *fkbb* and *fkbl* recombinant
genes of the invention can be used in heterologous hosts to produce compounds such as
FK-520 or, in conjunction with other PKS or NRPS genes, to produce known or novel
5 polyketides and non-ribosomal peptides.

The present invention also provides recombinant DNA compounds that encode
the P450 oxidase and methyltransferase genes involved in the biosynthesis of FK-520.
Figure 2 shows the various sites on the FK-520 polyketide core structure at which these
enzymes act. By providing these genes in recombinant form, the present invention
10 provides recombinant host cells that can produce FK-520. This is accomplished by
introducing the recombinant PKS, P450 oxidase, and methyltransferase genes into a
heterologous host cell. In a preferred embodiment, the heterologous host cell is
Streptomyces coelicolor CH999 or *Streptomyces lividans* K4-114, as described in U.S.
Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar.
15 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by
reference. In addition, by providing recombinant host cells that express only a subset of
these genes, the present invention provides methods for making FK-520 precursor
compounds not readily obtainable by other means.

In a related aspect, the present invention provides recombinant DNA compounds
20 and vectors that are useful in generating, by homologous recombination, recombinant
host cells that produce FK-520 precursor compounds. In this aspect of the invention, a
native host cell that produces FK-520 is transformed with a vector (such as an SCP2*
derived vector for *Streptomyces* host cells) that encodes one or more disrupted genes
(i.e., a hydroxylase, a methyltransferase, or both) or merely flanking regions from those
25 genes. When the vector integrates by homologous recombination, the native, functional
gene is deleted or replaced by the non-functional recombinant gene, and the resulting
host cell thus produces an FK-520 precursor. Such host cells can also be complemented
by introduction of a modified form of the deleted or mutated non-functional gene to
produce a novel compound.

30 In one important embodiment, the present invention provides a hybrid PKS and
the corresponding recombinant DNA compounds that encode those hybrid PKS
enzymes. For purposes of the present invention a hybrid PKS is a recombinant PKS that
comprises all or part of one or more modules and thioesterase/cyclase domain of a first
PKS and all or part of one or more modules, loading module, and thioesterase/cyclase

domain of a second PKS. In one preferred embodiment, the first PKS is all or part of the FK-520 PKS, and the second PKS is only a portion or all of a non-FK-520 PKS.

One example of the preferred embodiment is an FK-520 PKS in which the AT domain of module 8, which specifies a hydroxymalonyl CoA and from which the C-13 methoxy group of FK-520 is derived, is replaced by an AT domain that specifies a malonyl, methylmalonyl, or ethylmalonyl CoA. Examples of such replacement AT domains include the AT domains from modules 3, 12, and 13 of the rapamycin PKS and from modules 1 and 2 of the erythromycin PKS. Such replacements, conducted at the level of the gene for the PKS, are illustrated in the examples below. Another illustrative example of such a hybrid PKS includes an FK-520 PKS in which the natural loading module has been replaced with a loading module of another PKS. Another example of such a hybrid PKS is an FK-520 PKS in which the AT domain of module three is replaced with an AT domain that binds methylmalonyl CoA.

In another preferred embodiment, the first PKS is most but not all of a non-FK-520 PKS, and the second PKS is only a portion or all of the FK-520 PKS. An illustrative example of such a hybrid PKS includes an erythromycin PKS in which an AT specific for methylmalonyl CoA is replaced with an AT from the FK-520 PKS specific for malonyl CoA.

Those of skill in the art will recognize that all or part of either the first or second PKS in a hybrid PKS of the invention need not be isolated from a naturally occurring source. For example, only a small portion of an AT domain determines its specificity. See U.S. provisional patent application Serial No. 60/091,526, incorporated herein by reference. The state of the art in DNA synthesis allows the artisan to construct *de novo* DNA compounds of size sufficient to construct a useful portion of a PKS module or domain. For purposes of the present invention, such synthetic DNA compounds are deemed to be a portion of a PKS.

Thus, the hybrid modules of the invention are incorporated into a PKS to provide a hybrid PKS of the invention. A hybrid PKS of the invention can result not only:

(i) from fusions of heterologous domain (where heterologous means the domains in that module are from at least two different naturally occurring modules) coding sequences to produce a hybrid module coding sequence contained in a PKS gene whose product is incorporated into a PKS, but also:

(ii) from fusions of heterologous module (where heterologous module means two modules are adjacent to one another that are not adjacent to one another in naturally

occurring PKS enzymes) coding sequences to produce a hybrid coding sequence contained in a PKS gene whose product is incorporated into a PKS,

- (iii) from expression of one or more FK-520 PKS genes with one or more non-FK-520 PKS genes, including both naturally occurring and recombinant non-FK-520
- 5 PKS genes, and
- (iv) from combinations of the foregoing.

Various hybrid PKSs of the invention illustrating these various alternatives are described herein.

- Examples of the production of a hybrid PKS by co-expression of PKS genes from
- 10 the FK-520 PKS and another non-FK-520 PKS include hybrid PKS enzymes produced by coexpression of FK-520 and rapamycin PKS genes. Preferably, such hybrid PKS enzymes are produced in recombinant *Streptomyces* host cells that produce FK-520 or FK-506 but have been mutated to inactivate the gene whose function is to be replaced by the rapamycin PKS gene introduced to produce the hybrid PKS. Particular examples
- 15 include (i) replacement of the *fkfC* gene with the *rapB* gene; and (ii) replacement of the *fkfA* gene with the *rapC* gene. The latter hybrid PKS produces 13,15-didesmethoxy-FK-520, if the host cell is an FK-520 producing host cell, and 13,15-didesmethoxy-FK-506, if the host cell is an FK-506 producing host cell. The compounds produced by these hybrid PKS enzymes are immunosuppressants and neurotrophins but can be readily
- 20 modified to act only as neurotrophins, as described in Example 6, below.

- Other illustrative hybrid PKS enzymes of the invention are prepared by replacing the *fkfA* gene of an FK-520 or FK-506 producing host cell with a hybrid *fkfA* gene in which: (a) the extender module 8 through 10, inclusive, coding sequences have been replaced by the coding sequences for extender modules 12 to 14, inclusive, of the
- 25 rapamycin PKS; and (b) the module 8 coding sequences have been replaced by the module 8 coding sequence of the rifamycin PKS. When expressed with the other, naturally occurring FK-520 or FK-506 PKS genes and the genes of the modification enzymes, the resulting hybrid PKS enzymes produce, respectively, (a) 13-desmethoxy-FK-520 or 13-desmethoxy-FK-506; and (b) 13-desmethoxy-13-methyl-FK-520 or 13-
- 30 desmethoxy-13-methyl-FK-506. In a preferred embodiment, these recombinant PKS genes of the invention are introduced into the producing host cell by a vector such as pHU204, which is a plasmid pRM5 derivative that has the well-characterized SCP2* replicon, the *colE1* replicon, the *tsr* and *bla* resistance genes, and a *cos* site. This vector can be used to introduce the recombinant *fkfA* replacement gene in an FK-520 or FK-
- 35 506 producing host cell (or a host cell derived therefrom in which the endogenous *fkfA*

gene has either been rendered inactive by mutation, deletion or homologous recombination with the gene that replaces it) to produce the desired hybrid PKS.

In constructing hybrid PKSs of the invention, certain general methods may be helpful. For example, it is often beneficial to retain the framework of the module to be altered to make the hybrid PKS. Thus, if one desires to add DH and ER functionalities to a module, it is often preferred to replace the KR domain of the original module with a KR, DH, and ER domain-containing segment from another module, instead of merely inserting DH and ER domains. One can alter the stereochemical specificity of a module by replacement of the KS domain with a KS domain from a module that specifies a different stereochemistry. See Lau *et al.*, 1999, "Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units," *Biochemistry* 38(5):1643-1651, incorporated herein by reference. Stereochemistry can also be changed by changing the KR domain. Also, one can alter the specificity of an AT domain by changing only a small segment of the domain. See Lau *et al.*, *supra*. One can also take advantage of known linker regions in PKS proteins to link modules from two different PKSs to create a hybrid PKS. See Gokhale *et al.*, 16 Apr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases," *Science* 284: 482-485, incorporated herein by reference.

The following Table lists references describing illustrative PKS genes and corresponding enzymes that can be utilized in the construction of the recombinant PKSs and the corresponding DNA compounds that encode them of the invention. Also presented are various references describing tailoring enzymes and corresponding genes that can be employed in accordance with the methods of the present invention.†

Avermectin

U.S. Pat. No. 5,252,474 to Merck.

MacNeil *et al.*, 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

MacNeil *et al.*, 1992, *Gene* 115: 119-125, Complex Organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase.

Ikeda *et al.*, Aug. 1999, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. USA* 96: 9509-9514.

Candicidin (FR008)

Hu *et al.*, 1994, *Mol. Microbiol.* 14: 163-172.

Epothilone

U.S. Pat. App. Serial No. 60/130,560, filed 22 April 1999.

Erythromycin

5 PCT Pub. No. 93/13663 to Abbott.

US Pat. No. 5,824,513 to Abbott.

Donadio *et al.*, 1991, *Science* 252:675-9.

Cortes *et al.*, 8 Nov. 1990, *Nature* 348:176-8, An unusually large
multifunctional polypeptide in the erythromycin producing polyketide synthase of
10 *Saccharopolyspora erythraea*.

Glycosylation Enzymes

PCT Pat. App. Pub. No. 97/23630 to Abbott.

FK-506

Motamedi *et al.*, 1998, The biosynthetic gene cluster for the macrolactone ring of
15 the immunosuppressant FK-506, *Eur. J. biochem.* 256: 528-534.

Motamedi *et al.*, 1997, Structural organization of a multifunctional polyketide
synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506, *Eur.*
J. Biochem. 244: 74-80.

Methyltransferase

20 US 5,264,355, issued 23 Nov. 1993, Methylating enzyme from
Streptomyces MA6858. 31-O-desmethyl-FK-506 methyltransferase.

Motamedi *et al.*, 1996, Characterization of methyltransferase and
hydroxylase genes involved in the biosynthesis of the immunosuppressants FK-506 and
FK-520, *J. Bacteriol.* 178: 5243-5248.

25 *Streptomyces hygroscopicus*

U.S. patent application Serial No. 09/154,083, filed 16 Sep. 1998.

Lovastatin

U.S. Pat. No. 5,744,350 to Merck.

Narbomycin

30 U.S. patent application Serial No. 60/107,093, filed 5 Nov. 1998, and Serial No.
60/120,254, filed 16 Feb. 1999.

Nemadectin

MacNeil *et al.*, 1993, *supra*.

Niddamycin

Kakavas *et al.*, 1997, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol.* 179: 7515-7522.

Oleandomycin

Swan *et al.*, 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet.* 242: 358-362.

U.S. patent application Serial No. 60/120,254, filed 16 Feb. 1999.

Olano *et al.*, 1998, Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet.* 259(3): 299-308.

Picromycin

PCT patent application US99/15047, filed 2 Jul. 1999.

Xue *et al.*, 1998, Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the *pikC*-encoded cytochrome P450 in *Streptomyces venezuelae*, *Chemistry & Biology* 5(11): 661-667.

Xue *et al.*, Oct. 1998, A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity, *Proc. Natl. Acad. Sci. USA* 95: 12111-12116.

Platenolide

EP Pat. App. Pub. No. 791,656 to Lilly.

Rapamycin

Schwecke *et al.*, Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA* 92:7839-7843.

Aparicio *et al.*, 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, *Gene* 169: 9-16.

Rifamycin

August *et al.*, 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amiclatopsis mediterranei* S669, *Chemistry & Biology*, 5(2): 69-79.

Sorangium PKS

U.S. patent application Serial No. 09/144,085, filed 31 Aug. 1998.

Soraphen

U.S. Pat. No. 5,716,849 to Novartis.

Schupp *et al.*, 1995, *J. Bacteriology* 177: 3673-3679. A *Sorangium cellulosum* (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

5 **Spiramycin**

U.S. Pat. No. 5,098,837 to Lilly.

Activator Gene

U.S. Pat. No. 5,514,544 to Lilly.

Tylosin

10 EP Pub. No. 791,655 to Lilly.

U.S. Pat. No. 5,876,991 to Lilly.

Kuhstoss *et al.*, 1996, *Gene* 183:231-6., Production of a novel polyketide through the construction of a hybrid polyketide synthase.

Tailoring enzymes

15 Merson-Davies and Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355. Analysis of five tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

As the above Table illustrates, there are a wide variety of polyketide synthase genes that serve as readily available sources of DNA and sequence information for use in
20 constructing the hybrid PKS-encoding DNA compounds of the invention. Methods for constructing hybrid PKS-encoding DNA compounds are described without reference to the FK-520 PKS in PCT patent publication No. 98/51695; U.S. Patent Nos. 5,672,491 and 5,712,146 and U.S. patent application Serial Nos. 09/073,538, filed 6 May 1998, and 09/141,908, filed 28 Aug 1998, each of which is incorporated herein by reference.

25 The hybrid PKS-encoding DNA compounds of the invention can be and often are hybrids of more than two PKS genes. Moreover, there are often two or more modules in the hybrid PKS in which all or part of the module is derived from a second (or third) PKS. Thus, as one illustrative example, the present invention provides a hybrid FK-520 PKS that contains the naturally occurring loading module and FkbP as well as modules
30 one, two, four, six, seven, and eight, nine, and ten of the FK-520 PKS and further contains hybrid or heterologous modules three and five. Hybrid or heterologous module three contains an AT domain that is specific of methylmalonyl CoA and can be derived for example, from the erythromycin or rapamycin PKS genes. Hybrid or heterologous module five contains an AT domain that is specific for malonyl CoA and can be derived
35 for example, from the picromycin or rapamycin PKS genes.

While an important embodiment of the present invention relates to hybrid PKS enzymes and corresponding genes, the present invention also provides recombinant FK-520 PKS genes in which there is no second PKS gene sequence present but which differ from the FK-520 PKS gene by one or more deletions. The deletions can encompass one or more modules and/or can be limited to a partial deletion within one or more modules. When a deletion encompasses an entire module, the resulting FK-520 derivative is at least two carbons shorter than the gene from which it was derived. When a deletion is within a module, the deletion typically encompasses a KR, DH, or ER domain, or both DH and ER domains, or both KR and DH domains, or all three KR, DH, and ER domains.

To construct a hybrid PKS or FK-520 derivative PKS gene of the invention, one can employ a technique, described in PCT Pub. No. 98/27203 and U.S. patent application Serial No. 08/989,332, filed 11 Dec. 1997, each of which is incorporated herein by reference, in which the large PKS gene is divided into two or more, typically three, segments, and each segment is placed on a separate expression vector. In this manner, each of the segments of the gene can be altered, and various altered segments can be combined in a single host cell to provide a recombinant PKS gene of the invention. This technique makes more efficient the construction of large libraries of recombinant PKS genes, vectors for expressing those genes, and host cells comprising those vectors.

Thus, in one important embodiment, the recombinant DNA compounds of the invention are expression vectors. As used herein, the term expression vector refers to any nucleic acid that can be introduced into a host cell or cell-free transcription and translation medium. An expression vector can be maintained stably or transiently in a cell, whether as part of the chromosomal or other DNA in the cell or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a gene that serves to produce RNA that is translated into a polypeptide in the cell or cell extract. Furthermore, expression vectors typically contain additional functional elements, such as resistance-conferring genes to act as selectable markers.

The various components of an expression vector can vary widely, depending on the intended use of the vector. In particular, the components depend on the host cell(s) in which the vector will be used or is intended to function. Vector components for expression and maintenance of vectors in *E. coli* are widely known and commercially available, as are vector components for other commonly used organisms, such as yeast cells and *Streptomyces* cells.

In a preferred embodiment, the expression vectors of the invention are used to construct recombinant *Streptomyces* host cells that express a recombinant PKS of the invention. Preferred *Streptomyces* host cell/vector combinations of the invention include *S. coelicolor* CH999 and *S. lividans* K4-114 host cells, which do not produce
5 actinorhodin, and expression vectors derived from the pRM1 and pRM5 vectors, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference.

The present invention provides a wide variety of expression vectors for use in
10 *Streptomyces*. For replicating vectors, the origin of replication can be, for example and without limitation, a low copy number vector, such as SCP2* (see Hopwood *et al.*, *Genetic Manipulation of Streptomyces: A Laboratory manual* (The John Innes Foundation, Norwich, U.K., 1985); Lydiate *et al.*, 1985, *Gene* 35: 223-235; and Kieser and Melton, 1988, *Gene* 65: 83-91, each of which is incorporated herein by reference),
15 SLP1.2 (Thompson *et al.*, 1982, *Gene* 20: 51-62, incorporated herein by reference), and SG5(ts) (Muth *et al.*, 1989, *Mol. Gen. Genet.* 219: 341-348, and Bierman *et al.*, 1992, *Gene* 116: 43-49, each of which is incorporated herein by reference), or a high copy number vector, such as pIJ101 and pJV1 (see Katz *et al.*, 1983, *J. Gen. Microbiol.* 129: 2703-2714; Vara *et al.*, 1989, *J. Bacteriol.* 171: 5782-5781; and Servin-Gonzalez, 1993,
20 *Plasmid* 30: 131-140, each of which is incorporated herein by reference). Generally, however, high copy number vectors are not preferred for expression of genes contained on large segments of DNA. For non-replicating and integrating vectors, it is useful to include at least an *E. coli* origin of replication, such as from pUC, p1P, p1I, and pBR. For phage based vectors, the phages phiC31 and KC515 can be employed (see Hopwood
25 *et al.*, *supra*).

Typically, the expression vector will comprise one or more marker genes by which host cells containing the vector can be identified and/or selected. Useful antibiotic resistance conferring genes for use in *Streptomyces* host cells include the *ermE* (confers resistance to erythromycin and other macrolides and lincomycin), *tsr* (confers resistance
30 to thiostrepton), *aadA* (confers resistance to spectinomycin and streptomycin), *aacC4* (confers resistance to apramycin, kanamycin, gentamicin, geneticin (G418), and neomycin), *hyg* (confers resistance to hygromycin), and *vph* (confers resistance to viomycin) resistance conferring genes.

The recombinant PKS gene on the vector will be under the control of a promoter,
35 typically with an attendant ribosome binding site sequence. The present invention

provides the endogenous promoters of the FK-520 PKS and related biosynthetic genes in recombinant form, and these promoters are preferred for use in the native hosts and in heterologous hosts in which the promoters function. A preferred promoter of the invention is the *fkfO* gene promoter, comprised in a sequence of about 270 bp between the start of the open reading frames of the *fkfO* and *fkfB* genes. The *fkfO* promoter is believed to be bi-directional in that it promotes transcription of the genes *fkfO*, *fkfP*, and *fkfA* in one direction and *fkfB*, *fkfC*, and *fkfL* in the other. Thus, in one aspect, the present invention provides a recombinant expression vector comprising the promoter of the *fkfO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkfO*. In a preferred embodiment the transcribed gene is an FK-520 PKS gene. In another preferred embodiment, the transcribed gene is a gene that encodes a protein comprised in a hybrid PKS.

Heterologous promoters can also be employed and are preferred for use in host cells in which the endogenous FK-520 PKS gene promoters do not function or function poorly. A preferred heterologous promoter is the *actI* promoter and its attendant activator gene *actII-ORF4*, which is provided in the pRM1 and pRM5 expression vectors, *supra*. This promoter is activated in the stationary phase of growth when secondary metabolites are normally synthesized. Other useful *Streptomyces* promoters include without limitation those from the *ermE* gene and the *melC1* gene, which act constitutively, and the *tipA* gene and the *merA* gene, which can be induced at any growth stage. In addition, the T7 RNA polymerase system has been transferred to *Streptomyces* and can be employed in the vectors and host cells of the invention. In this system, the coding sequence for the T7 RNA polymerase is inserted into a neutral site of the chromosome or in a vector under the control of the inducible *merA* promoter, and the gene of interest is placed under the control of the T7 promoter. As noted above, one or more activator genes can also be employed to enhance the activity of a promoter. Activator genes in addition to the *actII-ORF4* gene discussed above include *dnrI*, *redD*, and *ptpA* genes (see U.S. patent application Serial No. 09/181,833, *supra*) to activate promoters under their control.

In addition to providing recombinant DNA compounds that encode the FK-520 PKS, the present invention also provides DNA compounds that encode the ethylmalonyl CoA and 2-hydroxymalonyl CoA utilized in the synthesis of FK-520. Thus, the present invention also provides recombinant host cells that express the genes required for the biosynthesis of ethylmalonyl CoA and 2-hydroxymalonyl CoA. Figures 3 and 4 show the

location of these genes on the cosmids of the invention and the biosynthetic pathway that produces ethylmalonyl CoA.

For 2-hydroxymalonyl CoA biosynthesis, the *fkbH*, *fkbl*, *fkbl*, and *fkbl* genes are sufficient to confer this ability on *Streptomyces* host cells. For conversion of 2-hydroxymalonyl to 2-methoxymalonyl, the *fkbl* gene is also employed. While the complete coding sequence for *fkbl* is provided on the cosmids of the invention, the sequence for this gene provided herein may be missing a T residue, based on a comparison made with a similar gene cloned from the ansamitocin gene cluster by Dr. H. Floss. Where the sequence herein shows one T, there may be two, resulting in an extension of the *fkbl* reading frame to encode the amino acid sequence:

MTIVKCLVWDLNLTWRGTVLEDDEVVLTDEIREVITTLDDRGLQAVASKNDH
DLAWERLERLGVAEYFVLARIGWGPKSQSVREIATELNFAPTTIAFIDDQPAERA
EVAFHLPEVRCYPAEQAATLLSLPEFSPPVSTVDSRRRLMYQAGFARDQAREA
YSGPDEDFLRSLDLSMTIAPAGEEELSRVEELTLRTSQMNATGVHYSDADLRAL
LTDPAHEVLVVTMGDRFGPHGAVGILLEKKPSTWHLKLLATSCRVVVSFGAGAT
ILNWLTDQGARAGAHLVADFRRTDRNRMMEIAYRFAGFADSDCPCVSEVAGAS
AAGVERLHLEPSARPAPTTTLTAADIAPVTVSAAG.

For ethylmalonyl CoA biosynthesis, one requires only a crotonyl CoA reductase, which can be supplied by the host cell but can also be supplied by recombinant expression of the *fkbl* gene of the present invention. To increase yield of ethylmalonyl CoA, one can also express the *fkbl* and *fkbl* genes as well. While such production can be achieved using only the recombinant genes above, one can also achieve such production by placing into the recombinant host cell a large segment of the DNA provided by the cosmids of the invention. Thus, for 2-hydroxymalonyl and 2-methoxymalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the left side of the FK-520 PKS genes shown in Figure 1. For ethylmalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the right side of the FK-520 PKS genes shown in Figure 1 or, alternatively, both the right and left segments of DNA.

The recombinant DNA expression vectors that encode these genes can be used to construct recombinant host cells that can make these important polyketide building blocks from cells that otherwise are unable to produce them. For example, *Streptomyces coelicolor* and *Streptomyces lividans* do not synthesize ethylmalonyl CoA or 2-hydroxymalonyl CoA. The invention provides methods and vectors for constructing recombinant *Streptomyces coelicolor* and *Streptomyces lividans* that are able to

synthesize either or both ethylmalonyl CoA and 2-hydroxymalonyl CoA. These host cells are thus able to make polyketides, those requiring these substrates, that cannot otherwise be made in such cells.

- In a preferred embodiment, the present invention provides recombinant
- 5 *Streptomyces* host cells, such as *S. coelicolor* and *S. lividans*, that have been transformed with a recombinant vector of the invention that codes for the expression of the ethylmalonyl CoA biosynthetic genes. The resulting host cells produce ethylmalonyl CoA and so are preferred host cells for the production of polyketides produced by PKS enzymes that comprise one or more AT domains specific for ethylmalonyl CoA.
- 10 Illustrative PKS enzymes of this type include the FK-520 PKS and a recombinant PKS in which one or more AT domains is specific for ethylmalonyl CoA.

- In a related embodiment, the present invention provides *Streptomyces* host cells in which one or more of the ethylmalonyl or 2-hydroxymalonyl biosynthetic genes have been deleted by homologous recombination or rendered inactive by mutation. For
- 15 example, deletion or inactivation of the *fkfG* gene can prevent formation of the methoxyl groups at C-13 and C-15 of FK-520 (or, in the corresponding FK-506 producing cell, FK-506), leading to the production of 13,15-didesmethoxy-13,15-dihydroxy-FK-520 (or, in the corresponding FK-506 producing cell, 13,15-didesmethoxy-13,15-dihydroxy-FK-506). If the *fkfG* gene product acts on 2-hydroxymalonyl and the resulting 2-
- 20 methoxymalonyl substrate is required for incorporation by the PKS, the AT domains of modules 7 and 8 may bind malonyl CoA and methylmalonyl CoA. Such incorporation results in the production of a mixture of polyketides in which the methoxy groups at C-13 and C-15 of FK-520 (or FK-506) are replaced by either hydrogen or methyl.

- This possibility of non-specific binding results from the construction of a hybrid
- 25 PKS of the invention in which the AT domain of module 8 of the FK-520 PKS replaced the AT domain of module 6 of DEBS. The resulting PKS produced, in *Streptomyces lividans*, 6-dEB and 2-desmethyl-6-dEB, indicating that the AT domain of module 8 of the FK-520 PKS could bind malonyl CoA and methylmalonyl CoA substrates. Thus, one could possibly also prepare the 13,15-didesmethoxy-FK-520 and corresponding FK-506
- 30 compounds of the invention by deleting or otherwise inactivating one or more or all of the genes required for 2-hydroxymalonyl CoA biosynthesis, i.e., the *fkfH*, *fkfI*, *fkfJ*, and *fkfK* genes. In any event, the deletion or inactivation of one or more biosynthetic genes required for ethylmalonyl and/or 2-hydroxymalonyl production prevents the formation of polyketides requiring ethylmalonyl and/or 2-hydroxymalonyl for biosynthesis, and the

resulting host cells are thus preferred for production of polyketides that do not require the same.

The host cells of the invention can be grown and fermented under conditions known in the art for other purposes to produce the compounds of the invention. See, e.g.,
5 U.S. Patent Nos. 5,194,378; 5,116,756; and 5,494,820, incorporated herein by reference, for suitable fermentation processes. The compounds of the invention can be isolated from the fermentation broths of these cultured cells and purified by standard procedures. Preferred compounds of the invention include the following compounds: 13-desmethoxy-FK-506; 13-desmethoxy-FK-520; 13,15-didesmethoxy-FK-506; 13,15-
10 didesmethoxy-FK-520; 13-desmethoxy-18-hydroxy-FK-506; 13-desmethoxy-18-hydroxy-FK-520; 13,15-didesmethoxy-18-hydroxy-FK-506; and 13,15-didesmethoxy-18-hydroxy-FK-520. These compounds can be further modified as described for tacrolimus and FK-520 in U.S. Patent Nos. 5,225,403; 5,189,042; 5,164,495; 5,068,323; 4,980,466; and 4,920,218, incorporated herein by reference.

15 Other compounds of the invention are shown in Figure 8, Parts A and B. In Figure 8, Part A, illustrative C-32-substituted compounds of the invention are shown in two columns under the heading R. The substituted compounds are preferred for topical administration and are applied to the dermis for treatment of conditions such as psoriasis. In Figure 8, Part B, illustrative reaction schemes for making the compounds shown in
20 Figure 8, Part A, are provided. In the upper scheme in Figure 8, Part B, the C-32 substitution is a tetrazole moiety, illustrative of the groups shown in the left column under R in Figure 8, Part A. In the lower scheme in Figure 8, Part B, the C-32 substitution is a disubstituted amino group, where R₃ and R₄ can be any group similar to the illustrative groups shown attached to the amine in the right column under R in Figure
25 8, Part A. While Figure 8 shows the C-32-substituted compounds in which the C-15-methoxy is present, the invention includes these C-32-substituted compounds in which C-15 is ethyl, methyl, or hydrogen. Also, while C-21 is shown as substituted with ethyl or allyl, the compounds of the invention includes the C-32-substituted compounds in which C-21 is substituted with hydrogen or methyl.

30 To make these C-32-substituted compounds, Figure 8, Part B, provides illustrative reaction schemes. Thus, a selective reaction of the starting compound (see Figure 8, Part B, for an illustrative starting compound) with trifluoromethanesulfonic anhydride in the presence of a base yields the C-32 O-triflate derivative, as shown in the upper scheme of Figure 8, Part B. Displacement of the triflate with 1H-tetrazole or
35 triazole derivatives provides the C-32 tetrazole or teiazole derivative. As shown in the

lower scheme of Figure 8, Part B, reacting the starting compound with p-nitrophenylchloroformate yields the corresponding carbonate, which, upon displacement with an amino compound, provides the corresponding carbamate derivative.

The compounds can be readily formulated to provide the pharmaceutical compositions of the invention. The pharmaceutical compositions of the invention can be used in the form of a pharmaceutical preparation, for example, in solid, semisolid, or liquid form. This preparation contains one or more of the compounds of the invention as an active ingredient in admixture with an organic or inorganic carrier or excipient suitable for external, enteral, or parenteral application. The active ingredient may be compounded, for example, with the usual non-toxic, pharmaceutically acceptable carriers for tablets, pellets, capsules, suppositories, solutions, emulsions, suspensions, and any other form suitable for use. Suitable formulation processes and compositions for the compounds of the present invention are described with respect to tacrolimus in U.S. Patent Nos. 5,939,427; 5,922,729; 5,385,907; 5,338,684; and 5,260,301, incorporated herein by reference. Many of the compounds of the invention contain one or more chiral centers, and all of the stereoisomers are included within the scope of the invention, as pure compounds as well as mixtures of stereoisomers. Thus the compounds of the invention may be supplied as a mixture of stereoisomers in any proportion.

The carriers which can be used include water, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate, talc, corn starch, keratin, colloidal silica, potato starch, urea, and other carriers suitable for use in manufacturing preparations, in solid, semi-solid, or liquified form. In addition, auxiliary stabilizing, thickening, and coloring agents and perfumes may be used. For example, the compounds of the invention may be utilized with hydroxypropyl methylcellulose essentially as described in U.S. Patent No. 4,916,138, incorporated herein by reference, or with a surfactant essentially as described in EPO patent publication No. 428,169, incorporated herein by reference.

Oral dosage forms may be prepared essentially as described by Hondo *et al.*, 1987, *Transplantation Proceedings XIX*, Supp. 6: 17-22, incorporated herein by reference. Dosage forms for external application may be prepared essentially as described in EPO patent publication No. 423,714, incorporated herein by reference. The active compound is included in the pharmaceutical composition in an amount sufficient to produce the desired effect upon the disease process or condition.

For the treatment of conditions and diseases relating to immunosuppression or neuronal damage, a compound of the invention may be administered orally, topically,

parenterally, by inhalation spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvant, and vehicles. The term parenteral, as used herein, includes subcutaneous injections, and intravenous, intramuscular, and intrasternal injection or infusion techniques.

5 Dosage levels of the compounds of the present invention are of the order from about 0.01 mg to about 50 mg per kilogram of body weight per day, preferably from about 0.1 mg to about 10 mg per kilogram of body weight per day. The dosage levels are useful in the treatment of the above-indicated conditions (from about 0.7 mg to about 3.5 mg per patient per day, assuming a 70 kg patient). In addition, the compounds of the
10 present invention may be administered on an intermittent basis, i.e., at semi-weekly, weekly, semi-monthly, or monthly intervals.

The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. For example, a formulation intended for oral
15 administration to humans may contain from 0.5 mg to 5 g of active agent compounded with an appropriate and convenient amount of carrier material, which may vary from about 5 percent to about 95 percent of the total composition. Dosage unit forms will generally contain from about 0.5 mg to about 500 mg of active ingredient. For external administration, the compounds of the invention can be formulated within the range of,
20 for example, 0.00001% to 60% by weight, preferably from 0.001% to 10% by weight, and most preferably from about 0.005% to 0.8% by weight. The compounds and compositions of the invention are useful in treating disease conditions using doses and administration schedules as described for tacrolimus in U.S. Patent Nos. 5,542,436; 5,365,948; 5,348,966; and 5,196,437, incorporated herein by reference. The compounds
25 of the invention can be used as single therapeutic agents or in combination with other therapeutic agents. Drugs that can be usefully combined with compounds of the invention include one or more immunosuppressant agents such as rapamycin, cyclosporin A, FK-506, or one or more neurotrophic agents.

It will be understood, however, that the specific dosage level for any particular
30 patient will depend on a variety of factors. These factors include the activity of the specific compound employed; the age, body weight, general health, sex, and diet of the subject; the time and route of administration and the rate of excretion of the drug; whether a drug combination is employed in the treatment; and the severity of the particular disease or condition for which therapy is sought.

5

Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-520

14

2

25

KC515). The ligation reactions contained 5 picomolar unphosphorylated linker DNA and 0.1 picomolar vector DNA, i.e., a 50-fold molar excess of linker to vector. The linker had the following sequence:

5 5'-CTAGTGGGCAGATCTGGCAGCT-3'
3'-ACCCGTCTAGACCG-5'

The resulting plasmid was designated pKOS60-27-1.

Next, a linker of the following sequence was ligated between the unique *Sph*I and *Afl*III sites of plasmid pKOS60-27-1 to introduce an *Nsi*I site at the 3' end of the module 8 cassette. The linker employed was:

10 5'-GGGATGCATGGC-3'
3'-GTACCCCTACGTACCGAATT-5'

The resulting plasmid was designated pKOS60-29-55.

To allow in-frame insertions of alternative AT domains, sites were engineered at the 5' end (*Avr* II or *Nhe* I) and 3' end (*Xho* I) of the AT domain using the polymerase chain reaction (PCR) as follows. Plasmid pKOS60-29-55 was used as a template for the PCR and sequence 5' to the AT domain was amplified with the primers *Spe*Bgl-fwd and either *Avr*-rev or *Nhe*-rev:

20 *Spe*Bgl-fwd 5'-CGACTCACTAGTGGGCAGATCTGG-3'
Avr-rev 5'-CACGCCTAGGCCGGTCGGTCTCGGGCCAC-3'
Nhe-rev 5'-GCGGCTAGCTGCTCGCCCATCGCGGGATGC-3'

The PCR included, in a 50 µl reaction, 5 µl of 10x *Pfu* polymerase buffer (Stratagene), 5 µl 10x z-dNTP mixture (2 mM dATP, 2 mM dCTP, 2 mM dTTP, 1 mM dGTP, 1 mM 7-deaza-GTP), 5 µl DMSO, 2 µl of each primer (10 µM), 1 µl of template DNA (0.1 µg/µl), and 1 µl of cloned *Pfu* polymerase (Stratagene). The PCR conditions were 95°C for 2 min., 25 cycles at 95°C for 30 sec., 60°C for 30 sec., and 72°C for 4 min., followed by 4 min. at 72°C and a hold at 0°C. The amplified DNA products and the Litmus vectors were cut with the appropriate restriction enzymes (*Bgl*II and *Avr*II or *Spe*I and *Nhe*I), and cloned into either pLitmus 28 or pLitmus38 (New England BioLabs), respectively, to generate the constructs designated pKOS60-37-4 and pKOS60-37-2, respectively.

Plasmid pKOS60-29-55 was again used as a template for PCR to amplify sequence 3' to the AT domain using the primers *Bsr*Xho-fwd and *Nsi*Afl-rev:

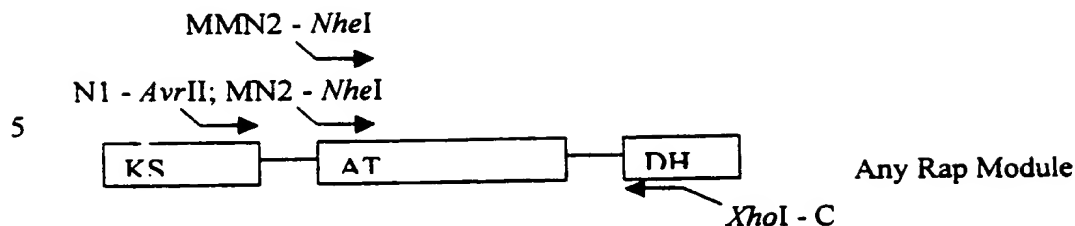
35 *Bsr*Xho-fwd 5'-GATGTACAGCTCGAGTCGGCACGCCCGGCCGCATC-3'
*Nsi*Afl-rev 5'-CGACTCACTTAAGCCATGCATCC-3'

PCR conditions were as described above. The PCR fragment was cut with *Bsr*GI and *Afl*III, gel isolated, and ligated into pKOS60-37-4 cut with *Asp*718 and *Afl*III and

inserted into pKOS60-37-2 cut with *Bsr*GI and *Afl*III, to give the plasmids pKOS60-39-1 and pKOS60-39-13, respectively. These two plasmids can be digested with *Avr*II and *Xho*I or *Nhe*I and *Xho*I, respectively, to insert heterologous AT domains specific for malonyl, methylmalonyl, ethylmalonyl, or other extender units.

- 5 Malonyl and methylmalonyl-specific AT domains were cloned from the rapamycin cluster using PCR amplification with a pair of primers that introduce an *Avr*II or *Nhe*I site at the 5' end and an *Xho*I site at the 3' end. The PCR conditions were as given above and the primer sequences were as follows:

- 10 RATN1 5'-ATCCTAGGCGGGCRGGYGTGTCGTCCTTCGG-3'
 (3' end of Rap KS sequence and universal for malonyl and methylmalonyl CoA),
 RATMN2 5'-ATGCTAGCCGCCGCGTTCCCCGTCTTCGCGCG-3'
 (Rap AT shorter version 5'- sequence and specific for malonyl CoA),
 RATMMN2 5'-ATGCTAGCGGATTCGTCGGTGGTGTTCGCCGA-3'
15 (Rap AT shorter version 5'- sequence and specific for methylmalonyl CoA), and
 RATC 5'-ATCTCGAGCCAGTASCGCTGGTGYTGGAAGG-3'
 (Rap DH 5'- sequence and universal for malonyl and methylmalonyl CoA).



Because of the high sequence similarity in each module of the rapamycin cluster, each primer was expected to prime any of the AT domains. PCR products representing ATs specific for malonyl or methylmalonyl extenders were identified by sequencing individual cloned PCR products. Sequencing also confirmed that the chosen clones contained no cloning artifacts. Examples of hybrid modules with the rapamycin AT12 and AT13 domains are shown in a separate figure.

The *AvrII*-*XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below. The AT of rap module 12 is specific for incorporation of malonyl units.

```

20 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
   I W Q L A E A L L T L V R E S T
   GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
   A A V L G H V G G E D I P A T A A
   GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150
25 F K D L G I D S L T A V Q L R N
   CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
   A L T E A T G V R L N A T A V F D
   TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250
   F P T P H V L A G K L G D E L T G
30 CACCCGCGCGCCCGTCTGTGCCCCGACCGCGGCCACGGCCGGTTCGCGACG 300
   T R A P V V P R T A A T A G A H
   ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350
   D E P L A I V G M A C R L P G G V
   GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
35 A S P E E L W H L V A S G T D A I
   CACGGAGTTCCCGACGGACCGCGGCTGGGACGTGACGCGATCTACGACC 450
   T E F P T D R G W D V D A I Y D
   CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
   P D P D A I G K T F V R H G G F L
40 ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCCGCGGA 550
   T G A T G F D A A F F G I S P R E
   GGCCCTCGCGATGGACCCGCAGCGGGTGTCTCTGGAGACGTCTGTGG 600
   A L A M D P Q Q R V L L E T S W
   AGGCGTTTCGAAAGCGCCGGCATACCCCGGACTCGACCCGCGGCAGCGAC 650
45 E A F E S A G I T P D S T R G S D
   ACCGGCGTGTTCGTGCGGCGCTTCTCTACGGTTACGGCACCGGTGCGGA 700
   T G V F V G A F S Y G Y G T G A D
   CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750
   T D G F G A T G S Q T S V L S G
50 GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800
   R L S Y F Y G L E G P A V T V D T
   GCGTGTTCGTGCTGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850
   A C S S S L V A L H Q A G Q S L R

```

CTCCGGCGAATGCTCGCTCGCCCTGGTCGGCGGCGTCACGGTGATGGCGT 900
S G E C S L A L V G G V T V M A
CTCCCGGCGGTTCTCGTGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950
S P G G F V E F S R Q R G L A P D
5 GGCCGGGCGAAGGCGTTCGGCGCGGGTGCGGACGGCACGAGCTTCGCCGA 1000
G R A K A F G A G A D G T S F A E
GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
G A G V L I V E R L S D A E R N
GTCACACCGTCTGGCGGTTCGTCCGTGGTTCGGCGGTCAACCAGGATGGT 1100
10 G H T V L A V V R G S A V N Q D G
GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTTCGCGAGGAGCGGGTGAT 1150
A S N G L S A P N G P S Q E R V I
CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGGCGGACGTGGACGCCG 1200
R Q A L A N A G L T P A D V D A
15 TCGAGGCCACGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250
V E A H G T G T R L G D P I E A Q
GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCTGCTGCTGGG 1300
A V L A T Y G Q E R A T P L L L G
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCGCCG 1350
20 S L K S N I G H A Q A A S G V A
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400
G I I K M V Q A L R H G E L P P T
CTGCACGCCGACGAGCCGTGCGCGCACGTTCGACTGGACGGCCGGCGCCGT 1450
L H A D E P S P H V D W T A G A V
25 CGAACTGCTGACGTGCGCCCGGCCGTGGCCCGAGACCGACCGGCCTAGGC 1500
E L L T S A R P W P E T D R P R
GGGCAGGCGTGTGCTCCTTCGGGATCAGTGGCACCAACGCCACGTCATC 1550
R A G V S S F G I S G T N A H V I
CTGGAAAGCGCACCCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG 1600
30 L E S A P P T Q P A D N A V I E R
GGCACCGGAGTGGGTGCCGTTGGTGATTTCCGGCCAGGACCCAGTCGGCTT 1650
A P E W V P L V I S A R T Q S A
TGACTGAGCACGAGGGCCGGTTGCGTGCGTATCTGGCGGCGTCGCCCGGG 1700
L T E H E G R L R A Y L A A S P G
35 GTGGATATGCGGGCTGTGGCATCGACGCTGGCGATGACACGGTCGGTGTT 1750
V D M R A V A S T L A M T R S V F
CGAGCACCGTGCCGTGCTGTTGGGAGATGACACCGTCACCGGCACCGCTG 1800
E H R A V L L G D D T V T G T A
TGTCTGACCCTCGGGCGGTGTTCTGCTCTTCCCGGGACAGGGGTTCGACGCT 1850
40 V S D P R A V F V F P G Q G S Q R
GCTGGCATGGGTGAGGAACCTGGCCGCCCGGTTCCCGCTCTTCGCGCGGAT 1900
A G M G E E L A A A F P V F A R I
CCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCCGATCTGGAGGTGAACG 1950
H Q Q V W D L L D V P D L E V N
45 AGACCGGTTACGCCCAGCCGGCCCTGTTTCGAATGCAGGTGGCTCTGTTC 2000
E T G Y A Q P A L F A M Q V A L F
GGGCTGCTGGAATCGTGGGGTGTACGACCGGACGCGGTGATCGGCCATTG 2050
G L L E S W G V R P D A V I G H S
GGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGTGTGGTTCGTTGGAGG 2100
50 V G E L A A A Y V S G V W S L E
ATGCCTGCACTTTGGTGTGCGCGCGGGCTCGTCTGATGCAGGCTCTGCCC 2150
D A C T L V S A R A R L M Q A L P
GCGGGTGGGGTGATGGTTCGTGTCCCGGTCTCGGAGGATGAGGCCCGGGC 2200
A G G V M V A V P V S E D E A R A
55 CGTGCTGGGTGAGGGTGTGGAGATCGCCCGGTCACGGCCCGTTCGTCGG 2250
V L G E G V E I A A V N G P S S
TGTTCTCTCCGGTGATGAGGCCGCCGTGCTGCAGGCCGCGGAGGGGCTG 2300
V V L S G D E A A V L Q A A E G L
GGGAAGTGGACGCGGCTGGCGACCGACCGGTTCCATTCCGCCCGTAT 2350
60 G K W T R L A T S H A F H S A R M
GGAACCCATGCTGGAGGAGTTCCGGGCGGTGCGCGAAGGCCTGACCTACC 2400
E P M L E E F R A V A E G L T Y
GGACGCCGAGGTCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAG 2450
R T P Q V S M A V G D Q V T T A E

TACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGC 2500
Y W V R Q V R D T V R F G E Q V A
CTCGTACGAGGACGCCGTGTTCTGTCGAGCTGGGTGCCGACCGGTCACTGG 2550
S Y E D A V F V E L G A D R S L
5 CCGGCCTGGTCGACGGTGTGCGGATGCTGCACGGCGACCACGAAATCCAG 2600
A R L V D G V A M L H G D H E I Q
GCCGCGATCGGCGCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCTGA 2650
A A I G A L A H L Y V N G V T V D
CTGGCCCGCGCTCCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTC 2700
10 W P A L L G D A P A T R V L D L
CGACATACGCCTTCCAGCACCAGCGCTACTGGCTCGAGTCGGCACGCCCG 2750
P T Y A F Q H Q R Y W L E S A R P
GCCGCGATCCGACGCGGGCCACCCCGTGGTGGCTCCGGTATCGCCCTCGC 2800
A A S D A G H P V L G S G I A L A
15 CCGGTGCGCGGGCCGGGTGTTACGGGTTCGTCGCGACCGGTGCGGACC 2850
G S P G R V F T G S V P T G A D
GCGCGGTGTTCTGTCGCGAGCTGGCGCTGGCCGCGCGGACGCGGTGCGAC 2900
R A V F V A E L A L A A A D A V D
TGCGCCACGGTCGAGCGGCTCGACATCGCCTCCGTGCGCGGGCGGGCGGG 2950
20 C A T V E R L D I A S V P G R P G
CCATGGCCGGGACGACCGTACAGACCTGGGTGCGACGAGCCGGCGGACGACG 3000
H G R T T V Q T W V D E P A D D
GCCGGCGCGGGTTCACCGTGCACACCCGACCGGCGACGCCCCGTGGACG 3050
G R R F T V H T R T G D A P W T
25 CTGCACGCGGAGGGGGTGTGCGCCCCATGGCACGGCCCTGCCCGATGC 3100
L H A E G V L R P H G T A L P D A
GGCCGACGCGGAGTGGCCCCCACCAGGGCGGGTGGCCGCGGACGGGCTGC 3150
A D A E W P P P G A V P A D G L
CGGGTGTGTGGCGCCGGGGGACCAGGTCTTCGCGGAGGCGGAGGTGGAC 3200
30 P G V W R R G D Q V F A E A E V D
GGACCGGACGGTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTC 3250
G P D G F V V H P D L L D A V F S
CGCGGTGCGCGACGGAAGCCGCCAGCCGGCGGATGGCGCGACCTGACGG 3300
A V G D G S R Q P A G W R D L T
35 TGCACGCGTCGGACGCCACCGTACTGCGCGCCTGCCTCACC CGGCGCACC 3350
V H A S D A T V L R A C L T R R T
GACGGAGCCATGGGATTGCGCGCCTTCGACGCGCGCGCCTGCCGGTACT 3400
D G A M G F A A F D G A G L P V L
CACCGCGGAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCG 3450
40 T A E A V T L R E V A S P S G S
AGGAGTCGGACGGCCTGCACCGTTGGAGTGGCTCGCGGTGCGCGGAGGCG 3500
E E S D G L H R L E W L A V A E A
GTCTACGACGGTGACCTGCCCCGAGGGACATGTCCTGATCACCGCCGCCA 3550
V Y D G D L P E G H V L I T A A H
45 CCCCAGCAGCCCCGAGGACATACCCACCCGCGCCACACCCGCGCCACCC 3600
P D D P E D I P T R A H T R A T
GCGTCCTGACCGCCCTGCAACACCACCTCACCACCACCGACACACCCCTC 3650
R V L T A L Q H H L T T T D H T L
ATCGTCCACACCACCGACCCCGCGCGGCCACCGTCACCGGCCTCAC 3700
50 I V H T T T D P A G A T V T G L T
CCGACCGCCCCAGAACGAACACCCCCACCGCATCCGCCTCATCGAAACCG 3750
R T A Q N E H P H R I R L I E T
ACCACCCCCACACCCCTCCCTGCCCCAACTCGCCACCCTCGACCAC 3800
D H P H T P L P L A Q L A T L D H
55 CCCCACCTCCGCTCACCCACCACCCCTCCACCACCCCCACCTCACCCC 3850
P H L R L T H H T L H H P H L T P
CCTCCACACCACCCACCCACCCACCCACCCCTCAACCCCGAACACG 3900
L H T T T P P T T T P L N P E H
CCATCATCATCACGGGCGGCTCCGGCACCCCTCGCCGGCATCCTCGCCCCG 3950
60 A I I I T G G S G T L A G I L A R
CACCTGAACACCCCCACACCTACCTCCTCCCGCACCCACCCCCCGA 4000
H L N H P H T Y L L S R T P P P D
CGCCACCCCGGCGACCCACCTCCCTGCGACGTGGGCGACCCCCACCAAC 4050
A T P G T H L P C D V G D P H Q

TCGCCACCACCTCACCACATCCCCAACCCTCACCGCCATCTTCCAC 4100
 L A T T L T H I P Q P L T A I F H
 ACCGCCGCCACCTCGACGACGGCATCTCCACGCCCTCACCCCGACCG 4150
 T A A T L D D G I L H A L T P D R
 5 CCTCACCACCGTCTCCACCCAAAGCCAACGCCGCTGGCACCTGCACC 4200
 L T T V L H P K A N A A W H L H
 ACCTCACCACCAACCCCTCACCACCTTCGTCTCTACTCCAGCGCC 4250
 H L T Q N Q P L T H F V L Y S S A
 GCCGCCGTCTCGGCAGCCCCGGACAAGGAACTACGCCGCCGCCAACGC 4300
 10 A A V L G S P G Q G N Y A A A N A
 CTTCTCGACGCCCTCGCCACCCACCGCCACACCTCGGCCAACCCGCCA 4350
 F L D A L A T H R H T L G Q P A
 CCTCCATCGCCTGGGGCATGTGGCACACCACAGCACCTCACCAGGACAA 4400
 T S I A W G M W H T T S T L T G Q
 15 CTCGACGACGCCGACCGGGACCGCATCCGCCGCGCGGTTTCTCCCGAT 4450
 L D D A D R D R I R R G G F L P I
 CACGGACGACGAGGGCATGGGGATGCAT
 T D D E G

- 20 The *AvrII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
 25 Q L A E A L L T L V R E S T
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
 A A V L G H V G G E D I P A T A A
 GTTCAAGGACCTCGGCATCGACTCGCTCACC CGGTCCAGCTGCGCAACG 150
 F K D L G I D S L T A V Q L R N
 30 CCCTCACCAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
 A L T E A T G V R L N A T A V F D
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAAGTACCGG 250
 F P T P H V L A G K L G D E L T G
 CACCCGCGCGCTCGTCTGCCCCGACCGCGGCCACGGCCCGGTGCGCACG 300
 35 T R A P V V P R T A A T A G A H
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGCTGCCCGCGGGGTC 350
 D E P L A I V G M A C R L P G G V
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
 A S P E E L W H L V A S G T D A I
 40 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTCGACGCGATCTACGACC 450
 T E F P T D R G W D V D A I Y D
 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
 P D P D A I G K T F V R H G G F L
 ACCGGCGGACAGGCTTCGACGCGGCTTCTTCGGCATCAGCCCGCGCGA 550
 45 T G A T G F D A A F F G I S P R E
 GGCCCTCGCGATGGACCCGACGAGCGGGTGCTCCTGGAGACGTCTGTGGG 600
 A L A M D P Q Q R V L L E T S W
 AGGCGTTGAAAGCGCCGGCATCACC CGGACTCGACCCGCGGCAGCGAC 650
 E A F E S A G I T P D S T R G S D
 50 ACCGGCGTGTTCGTGCGCGCCTTCTCTACGTTACGGCACCGGTGCGGA 700
 T G V F V G A F S Y G Y G T G A D
 CACCGACGGCTTCGGCGGACCGGCTCGCAGACAGTGTGCTCTCCGGCC 750
 T D G F G A T G S Q T S V L S G
 GGCTGTGCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800
 55 R L S Y F Y G L E G P A V T V D T
 GCGTGTTCGTGCTGCTGGTGGCGCTGCACCAGGCCGGCAGTCTGCTGCG 850
 A C S S S L V A L H Q A G Q S L R
 CTCCGGCGAATGCTCGCTCGCCCTGGTCCGGCGGCTCACGGTGATGGCGT 900
 S G E C S L A L V G G V T V M A
 60 CTCCCGCGGCTTCGTGGAGTTCTCCCGGACGCGCGGCTCGCGCCGGAC 950

S P G G F V E F S R Q R G L A P D
GGCCGGGCGAAGGCGTTCGGCGCGGGTGCAGGACGGCAGAGCTTCGCCGA 1000
G R A K A F G A G A D G T S F A E
GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
5 G A G V L I V E R L S D A E R N
GTCACACCGTCCCTGGCGGTTCGTCGGTTCGGCGGTCAACCAGGATGGT 1100
G H T V L A V V R G S A V N Q D G
GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTTCGAGGAGCGGGTGAT 1150
A S N G L S A P N G P S Q E R V I
10 CCGGCAGGCCCTGGCCAACGCCGGGTACCCCGCGGACGTGGACGCCG 1200
R Q A L A N A G L T P A D V D A
TCGAGGCCCCACGGCACCGGCACCGGCTGGGCGACCCCATCGAGGCACAG 1250
V E A H G T G T R L G D P I E A Q
GCGGTACTGGCCACTACGGACAGCGCGCCACCCCTGCTGCTGGG 1300
15 A V L A T Y G Q E R A T P L L L G
CTCGTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCGCCG 1350
S L K S N I G H A Q A A S G V A
GCATCATCAAGATGGTGCAGGCGCTCCGGCACGGGAGCTGCCGCCGACG 1400
G I I K M V Q A L R H G E L P P T
20 CTGACGCGGACGAGCCGTCGCCGACGTCGACTGGACGGCCGGCGCCGT 1450
L H A D E P S P H V D W T A G A V
CGAACTGCTGACGTCGGCCCCGGCGTGGCCCCGAGACCGACCGGCCTAGGC 1500
E L L T S A R P W P E T D R P R
GGGCGGGCGTGTGCTCCTTCGGAGTCAGCGGCACCAACGCCACGTCATC 1550
25 R A G V S S F G V S G T N A H V I
CTGGAGAGCGCACCCCCGCTCAGCCCCGGGAGGAGGCGCAGCCTGTTGA 1600
L E S A P P A Q P A E E A Q P V E
GACGCGGGTGGTGGCCTCGGATGTGCTGCCGCTGGTGATATCGGCCAAGA 1650
T P V V A S D V L P L V I S A K
30 CCCAGCCCCGCTGACCGAACACGAAGACCGGCTGCGCGCTACCTGGCG 1700
T Q P A L T E H E D R L R A Y L A
GCGTCGCCCCGGGGCGGATATACGGGCTGTGGCATCGACGCTGGCGGTGAC 1750
A S P G A D I R A V A S T L A V T
ACGGTCGGTGTTCGAGCACCGCGCCGTACTCCTTGGAGATGACACCGTCA 1800
35 R S V F E H R A V L L G D D T V
CCGGCACCGCGGTGACCGACCCAGGATCGTGTGTTGCTTTCCCGGGCAG 1850
T G T A V T D P R I V F V F P G Q
GGGTGGCAGTGCTGGGGATGGGCGAGTGCAGTGCAGGATTCGTCGGTGGT 1900
G W Q W L G M G S A L R D S S V V
40 GTTCGCGGAGCGGATGGCCGAGTGTGCGGCGGCGTTGCGCGAGTTCGTGG 1950
F A E R M A E C A A A L R E F V
ACTGGGATCTGTTACGGTTCTGGATGATCCGGCGGTGGTGGACCGGGTT 2000
D W D L F T V L D D P A V V D R V
GATGTGGTCCAGCCCGCTTCTGCGGATGATGGTTTCCCTGGCCGCGGT 2050
45 D V V Q P A S W A M M V S L A A V
GTGGCAGGCGGCGGTGTGCGGCGGATGCGGTGATCGGCCATTTCGACAG 2100
W Q A G V R P D A V I G H S Q
GTGAGATCGCCGAGCTTGTGTGGCGGGTGCAGTGTACTACGCGATGCC 2150
G E I A A A C V A G A V S L R D A
50 GCCCGGATCGTGACCTTGCAGCAGCCAGGCGATCGCCCGGGGCTGGCGGG 2200
A R I V T L R S Q A I A R G L A G
CCGGGGCGCGATGGCATCCGTCGCCCTGCCGCGCAGGATGTGAGCTGG 2250
R G A M A S V A L P A Q D V E L
TCGACGGGGCTGGATCGCCGCCACCAACGGGCGCCCTCCACCGTGATC 2300
55 V D G A W I A A H N G P A S T V I
GCGGGCACCCCGGAAGCGGTGACCATGTCCTACCGCTCATGAGGCACA 2350
A G T P E A V D H V L T A H E A Q
AGGGGTGCGGGTGCAGGCGGATCACCGTCGACTATGCCTCGCACACCCCGC 2400
G V R V R R I T V D Y A S H T P
60 ACGTCGAGCTGATCCGCGACGAATACTCGACATCACTAGCGACAGCAGC 2450
H V E L I R D E L L D I T S D S S
TCGACAGACCCCGCTCGTGCCGTGGCTGTGACCGTGGACGGCACCTGGGT 2500
S Q T P L V P W L S T V D G T W V
CGACAGCCCGCTGGACGGGGAGTACTGGTACCGGAACCTGCGTGAACCGG 2550

D S P L D G E Y W Y R N L R E P
 TCGGTTTCCACCCCGCCGTCAGCCAGTTGACAGGCCAGGGCGACACCGTG 2600
 V' G F H P A V S Q L Q A Q G D T V
 TTCGTCGAGGTGACGCCAGCCCGGTGTTGTTGACAGCGATGGACGACGA 2650
 5 F V E V S A S P V L L Q A M D D D
 TGTCGTCACGGTTGCCACGCTGCGTCGTGACGACGGCGACGCCACCCGGA 2700
 V V T V A T L R R D D G D A T R
 TGCTCACCGCCCTGGCACAGGCCTATGTCCACGGCGTCACCGTCGACTGG 2750
 M L T A L A Q A Y V H G V T V D W
 10 CCCGCCATCCTCGGCACCACCACAACCCGGTACTGGACCTTCCGACCTA 2800
 P A I L G T T T T R V L D L P T Y
 CGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGGCACGCCCGGCCGAT 2850
 A F Q H Q R Y W L E S A R P A A
 CCGACGCGGGGCCACCCCGTGCTGGGCTCCGGTATCGCCCTCGCCGGGTG 2900
 15 S D A G H P V L G S G I A L A G S
 CCGGGCCGGGTGTTACGGGTTCCGTGCCGACCGGTGCGGACCGCGCGGT 2950
 P G R V F T G S V P T G A D R A V
 GTTCGTCGCCGAGCTGGCGCTGGCCGCCGCGGACGCGGTGACTGCGCCA 3000
 F V A E L A L A A A D A V D C A
 20 CGGTGACGCGGCTCGACATCGCCTCCGTGCCCGGCCGGCCGGGCCATGGC 3050
 T V E K L D I A S V P G R P G H G
 CGGACGACCGTACAGACCTGGGTGACGAGCCGGCGGACGACGGCCGGCG 3100
 R T T V Q T W V D E P A D D G R R
 CCGGTTACCGTGACACCCGACCGGCGACGCCCCGTGGACGCTGCACG 3150
 25 R F T V H T R T G D A P W T L H
 CCGAGGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATGCGGCCGAC 3200
 A E G V L R P H G T A L P D A A D
 GCCGAGTGCCCCCACC GGCGCGGTGCCCGCGGACGGGCTGCCGGGTGT 3250
 A E W P P P G A V P A D G L P G V
 30 GTGGCGCCGGGGGGACCAGGTCTTCGCCGAGGCCGAGGTGGACGGACCGG 3300
 W R R G D Q V F A E A E V D G P
 ACGGTTTCGTGGTGCACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTC 3350
 D G F V V H P D L L D A V F S A V
 GGCGACGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGCACGC 3400
 35 G D G S R P A G W R D L T V H A
 GTCGGACGCCACCGTACTGCGCGCCTGCCTACCCGGCGCACCGACGGAG 3450
 S D A T V L R A C L T R R T D G
 CCATGGGATTGCGCGCCTTCGACGGCGCGGCCTGCCGGTACTCACCGCG 3500
 A M G F A A F D G A G L P V L T A
 40 GAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTC 3550
 E A V T L R E V A S P S G S E E S
 GGACGGCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACG 3600
 D G L H R L E W L A V A E A V Y
 ACGGTGACCTGCCCCAGGGACATGTCTGATCACCGCCGCCCCACCCCGAC 3650
 45 D G D L P E G H V L I T A A H P D
 GACCCCGAGGACATACCCACCCGCGCCACACCCGCGCCACCCGCGTCCT 3700
 D P E D I P T R A H T R A T R V L
 GACCGCCCTGCAACACCACCTCACCACCACCGACACACCTCATCGTCC 3750
 T A L Q H H L T T T D H T L I V
 50 ACACCACCACCGACCCCGCGCGCCACCGTCACCGGCCTCACCCGCACC 3800
 H T T T D P A G A T V T G L T R T
 GCCCAGAACGAACACCCCCACCGCATCCGCCTCATCGAAACCGACACCC 3850
 A Q N E H P H R I R L I E T D H P
 CCACACCCCCCTCCCCCTGGCCCAACTCGCCACCCCTCGACCACCCCCACC 3900
 55 H T P L A Q L A T L D H P H
 TCCGCCTCACCCACCACACCTCCACCACCCCCACCTCACCCCCCTCCAC 3950
 L R L T H H T L H H P H L T P L H
 ACCACCACCCACCCACCACCCCCCTCAACCCGAACACGCCATCAT 4000
 T T T P P T T T P L N P E H A I I
 60 CATCACGGGCGGCTCCGGCACCCCTCGCCGGCATCTCGCCCGCCACCTGA 4050
 I T G G S G T L A G I L A R H L
 ACCACCCACACCTACCTCCTCTCCCGCACCCACCCCCGACGCCACC 4100
 N H P H T Y L L S R T P P P D A T
 CCCGACCCACCTCCCCTGCGACGTGCGGACCCCCACCAACTCGCCAC 4150

P G T H L P C D V G D P H Q L A T
 CACCCTCACCACATCCCCAACCCCTCACCGCCATCTTCCACACCGCCG 4200
 T L T H I P Q P L T A I F H T A
 CCACCCTCGACGACGGCATCCTCCACGCCCTCACCCCCGACCGCCTCACC 4250
 5 A T L D D G I L H A L T P D R L T
 ACCGTCCTCCACCCCAAAGCCAACGCCGCTGGCACCTGCACCACCTCAC 4300
 T V L H P K A N A A W H L H H L T
 CCAAACCAACCCCTCACCACCTTCGTCTCTACTCCAGCGCCGCCGCCG 4350
 Q N Q P L T H F V L Y S S A A A
 10 TCCTCGGCAGCCCCGGACAAGGAACTACGCCGCCGCCAACGCCTTCCTC 4400
 V L G S P G Q G N Y A A A N A F L
 GAGCCCTCGCCACCCACCGCCACCCCTCGGCCAACCCGCCACCTCCAT 4450
 D A L A T H R H T L G Q P A T S I
 CGCCTGGGGCATGTGGCACACCACCAGCACCTCACCGGACAACCTCGACG 4500
 15 A W G M W H T T S T L T G Q L D
 ACGCCGACCGGGACCGCATCCGCCCGGGCGGTTTCTCCCGATCACGGAC 4550
 D A D R D R I R R G G F L P I T D
 GACGAGGGCATGGGGATGCAT
 D E G
 20

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS
 with the endogenous AT domain replaced by the AT domain of module 12 (specific for
 malonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid
 sequence shown below.

25 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
 Q L A E A L L T L V R E S T
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
 A A V L G H V G G E D I P A T A A
 GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150
 30 F K D L G I D S L T A V Q L R N
 CCCTCACCAGGCGACCGGTGTGCGGTGAACGCCACGGCGGTCTTCGAC 200
 A L T E A T G V R L N A T A V F D
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGGCAGCAACTGACCGG 250
 F P T P H V L A G K L G D E L T G
 35 CACCCGCGCGCCCGTCTGTGCCCCGACCGCGGCCACGGCCGCTGCGCACG 300
 T R A P V V P R T A A T A G A H
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGGCGGGTC 350
 D E P L A I V G M A C R L P G G V
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCAGCGCCAT 400
 40 A S P E E L W H L V A S G T D A I
 CACGGAGTTCCCGACGGACCGCGGTGGGACGTGACGCGATCTACGACC 450
 T E F P T D R G W D V D A I Y D
 CGGACCCCGACCGCATCGGCAAGACCTTCGTCCGGCAGGTTGGCTTCCTC 500
 P D P D A I G K T F V R H G G F L
 45 ACCGGCGCGACAGGCTTCGACGCGGCTTCTTCGGCATCAGCCCGCGCGA 550
 T G A T G F D A A F F G I S P R E
 GGCCCTCGCGATGGACCCGACGAGCGGGTGCTCCTGGAGACGTGCTGGG 600
 A L A M D P Q Q R V L L E T S W
 AGGCGTTCAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650
 50 E A F E S A G I T P D S T R G S D
 ACCGGCGTGTTCGTGCGCGCCTTCTCTACGGTTACGGCACCAGGTGCGGA 700
 T G V F V G A F S Y G Y G T G A D
 CACCGACGGCTTCGGCGCGACCGGCTCGCAGACCAGTGTGCTCTCCGGCC 750
 T D G F G A T G S Q T S V L S G
 55 GGCTGTGCTACTTCTACGGTCTGGAGGTCGGCGGTACGGTCGACACG 800
 R L S Y F Y G L E G P A V T V D T
 GCGTGTTCGTGCTGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850
 C S S L V A L H Q A G Q S L R
 CTCCGGCGAATGCTCGCTCGCCCTGGTCCGGCGGCTCACGGTGATGGCGT 900
 60 S G E C S L A L V G G V T V M A

CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950
S P G G F V E F S R Q R G L A P D
GGCCGGGCGAAGGCGTTCCGGCGGGTGCGGACGGCAGAGCTTCGCCGA 1000
G R A K A F G A G A D G T S F A E
5 GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
G A G V L I V E R L S D A E R N
GTCACACCGTCTCGGCGGTGTCGCGGTCAACCAGGATGGT 1100
G H T V L A V V R G S A V N Q D G
GCCTCCAACGCGCTGTGCGCGCGGAACGGGCGCTCGCAGGAGCGGGTGAT 1150
10 A S N G L S A P N G P S Q E R V I
CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGGCGGACGTGGACGCCG 1200
R Q A L A N A G L T P A D V D A
TCGAGGCCCCACGGCACCGGCACCGAGGCTGGGCGACCCCATCGAGGCACAG 1250
V E A H G T G T R L G D P I E A Q
15 GCGGTACTGGCCACTACGGACAGGAGCGGCCACCCCTGCTGCTGGG 1300
A V L A T Y G Q E R A T P L L L G
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCCGCCG 1350
S L K S N I G H A Q A A S G V A
GCATCATCAAGATGGTGCAGGCCCTCCGGCAGGGGAGCTGCCGCCGACG 1400
20 G I I K M V Q A L R H G E L P P T
CTGCACGCCGACGAGCCGTGCGCGCACGTGCGACTGGACGGCCGGCGCCGT 1450
L H A D E P S P H V D W T A G A V
CGAACTGCTGACGTGCGCCCGGCCGTGGCCCGAGACCGACCGGCCACGGC 1500
E L L T S A R P W P E T D R P R
25 GTGCCGCCGTCTCCTCGTTCCGGGGTGAGCGGCACCAACGCCACGTCATC 1550
R A A V S S F G V S G T N A H V I
CTGGAGGCCGACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGA 1600
L E A C P V T E T P A A S P S G D
CCTTCCCTGCTGGTGTGCGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650
30 L P L L V S A R S P E A L D E Q
TCCGCCGACTGCGCGCCTACCTGGACACCACCCCGGACGTGACCGGGTG 1700
I R R L R A Y L D T T P D V D R V
GCCGTGGCACAGACGCTGGCCCGGCGCACACTTCGCCACCGCGCCGT 1750
A V A Q T L A R R T H F A H R A V
35 GCTGCTCGGTGACACCGTCATCACACACCCCGCGGACCGGCCGACG 1800
L L G D T V I T T P P A D R P D
AACTCGTCTTCGTCTACTCCGGCCAGGGCACCCAGCATCCCGCGATGGGC 1850
E L V F V Y S G Q G T Q H P A M G
GAGCAGCTAGCCGCCGCGTTCCTCCCGTCTTCGCGCGGATCCATCAGCAGT 1900
40 E Q L A T A A F P V F A R I H Q Q V
GTGGGACCTGCTCGATGTGCCGATCTGGAGGTGAACGAGACCGGTTACG 1950
W D L L D V P D L E V N E T G Y
CCCAGCCGGCCCTGTTCCGAATGCAGGTGGCTCTGTTCCGGGCTGCTGGAA 2000
A Q P A L F A M Q V A L F G L L E
45 TCGTGGGGTGACGACCGGACGCGGTGATCGGCCATTCCGTGGGTGAGCT 2050
S W G V R P D A V I G H S V G E L
TGCGGCTGCGTATGTGTCCGGGGTGTGGTCTGGAGGATGCCTGCACTT 2100
A A A Y V S G V W S L E D A C T
TGGTGTGCGCGCGGGCTCGTCTGATGCAGGCTCTGCCCGCGGGTGGGGTG 2150
50 L V S A R A R L M Q A L P A G G V
ATGGTCTGCTGTCCCGGTCTCGGAGGATGAGGCCCGGGCCGTGCTGGGTGA 2200
M V A V P V S E D E A R A V L G E
GGGTGTGAGATCGCCGCGGTCAACGGCCCGTCTGCGGTGGTTCTCTCCG 2250
G V E I A A V N G P S S V V L S
55 GTGATGAGGCCGCGGTGCTGTCAGGCGCGGAGGGGCTGGGGAAGTGGACG 2300
G D E A A V L Q A A E G L G K W T
CGGCTGGCGACCGACCGGTTCCATTCCGCCCGTATGGAACCCATGCT 2350
R L A T S H A F H S A R M E P M L
GGAGGAGTTCGGGCGGTGCGCCGAAGGCCTGACCTACCGGACGCCGACG 2400
60 E F R A V A E G L T Y R T P Q
TCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGG 2450
V S M A V G D Q V T T A E Y W V R
CAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGA 2500
Q V R D T V R F G E Q V A S Y E D

CGCCGTGTTTCGTCGAGCTGGGTGCCGACCGGTCACTGGCCCCGCTGGTTCG 2550
A V F V E L G A D R S L A R L V
ACGGTGTGCGGATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGC 2600
D G V A M L H G D H E I Q A A I G
5 GCCCTGGCCCCACCTGTATGTCAACGGCGTCACGGTTCGACTGGCCCCGCGCT 2650
A L A H L Y V N G V T V D W P A L
CCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCT 2700
L G D A P A T R V L D L P T Y A
TCCAGCACCAGCGCTACTGGCTCGAGTCGGCACGCCCCGCGCATCCGAC 2750
10 F Q H Q R Y W L E S A R P A A S D
GCGGGCCACCCCGTGTCTGGGCTCCGGTATCGCCCTCGCCGGGTGCGCGGG 2800
A G H P V L G S G I A L A G S P G
CCGGTGTTCACGGGTTCGGTGCCGACCGGTGCGGACCGCGCGGTGTTCG 2850
R V F T G S V P T G A D R A V F
15 TCGCCGAGCTGGCGCTGGCCGCGCGGACGCGGTTCGACTGCGCCACGGTC 2900
V A E L A L A A A D A V D C A T V
GAGCGGCTCGACATCGCCTCCGTGCCCCGCGCGCGGCCATGGCCGGAC 2950
E R L D I A S V P G R P G H G R T
GACCGTACAGACCTGGGTTCGACGAGCCGGCGGACGACGCGCGCGCGGT 3000
20 T V Q T W V D E P A D D G R R R
TCACCGTGCACACCCGACCGCGACGCCCCGTGGACGCTGCACGCCGAG 3050
F T V H T R T G D A P W T L H A E
GGGGTGTCTCGCCCCCATGGCACGGCCCTGCCCGATGCGGCGGACGCCGA 3100
G V L R P H G T A L P D A A D A E
25 GTGGCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGCCGGGTGTGTGGC 3150
W P P P G A V P A D G L P G V W
GCCGGGGGACAGGTCTTCGCGGAGGCCGAGGTGGACGGACCGGACGGT 3200
R R G D Q V F A E A E V D G P D G
TTCGTGGTGCACCCGACCTGCTCGACGCGGTCTTCTCCGCGGTGCGCGA 3250
30 F V V H P D L L D A V F S A V G D
CGGAAGCCGCCAGCCGGCGCGGATGGCGCGACCTGACGGTGCACGCGTCGG 3300
G S R Q P A G W R D L T V H A S
ACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACCGACGGAGCCATG 3350
D A T V L R A C L T R R T D G A M
35 GGATTCCGCGCCTTCGACGGCGCGGCGCTGCCGGTACTCACCGCGGAGGC 3400
G F A A F D G A G L P V L T A E A
GGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTCGGACG 3450
V T L R E V A S P S G S E E S D
GCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACGACGGT 3500
40 G L H R L E A V A E A V Y D G
GACCTGCCCCGAGGACATGTCTGATCACCGCGCCACCCCGACGACCC 3550
D L P E G H V L I T A A H P D D P
CGAGGACATACCCACCCGCGCCACACCCGCGCCACCCGCGTCTTGACCG 3600
E D I P T R A H T R A T R V L T
45 CCCTGCAACACCACCTCACCACCACCGACCACCCCTCATCGTCCACACC 3650
A L Q H H L T T T D H T L I V H T
ACCACGACCCCGCGGCGCCACCGTACCGGCTCACCCGCACCGCCCA 3700
T T D P A G A T V T G L T R T A Q
GAACGAACACCCCAACCGCATCCGCTCATCGAAACCGACACCCCA 3750
50 N E H P H R I R L I E T D H P H
CCCCCTCCCCCTGGCCCAACTCGCCACCCTCGACCACCCCACTCCGC 3800
T P L P L A Q L A T L D H P H L R
CTCACCCACCAACCCCTCCACCAACCCCACTCACCCCTCCACACCAC 3850
L T H H T L H P H L T P L H T T
55 CACCCCAACCCACCAACCCCTCAACCCGAACACGCCATCATCATCA 3900
T P P T T T P L N P E H A I I I
CCGGCGGTCCGGCACCTCGCCGGCATCCTCGCCCGCCACCTGAACCAC 3950
T G G S G T L A G I L A R H L N H
CCCCACACCTACCTCCTCTCCCGCACCCACCCCGACGCCACCCCGG 4000
60 P H T Y L L S R T P P P D A T P G
CACCCACCTCCCCTGCGACGTGGGCGACCCCACTCGCCACCAACC 4050
T H L P C D V G D P H Q L A T T
TCACCCACATCCCCCAACCCCTCACCGCATCTTCCACACCGCGCCACC 4100
L T H I P Q P L T A I F H T A A T

CTCGACGACGGCATCCTCCACGCCCTCACCCCGACCGCTCACCACCGT 4150
 L D D G I L H A L T P D R L T T V
 CCTCCACCCCAAGCCAACGCCGCTGGCACCTGCACCACCTCACCCAAA 4200
 L H P K A N A A W H L H H L T Q
 5 ACCAACCCCTCACCACTTCGTCTCTACTCCAGCGCGCGCGCTCCTC 4250
 N Q P L T H F V L Y S S A A A V L
 GGCAGCCCGGACAAGGAACTACGCCGCGCCCAACGCCTTCCTCGACGC 4300
 G S P G Q G N Y A A A N A F L D A
 CCTCGCCACCCACCGCCACACCCTCGGCCAACCGCCACCTCCATCGCCT 4350
 10 L A T H R H T L G Q P A T S I A
 GGGGCATGTGGCACACCACCAGCACCTCACCGGACAACCTCGACGACGCC 4400
 W G N W H T T S T L T G Q L D D A
 GACCGGGACCGCATCCGCCGCGCGGTTTCTCCCGATCACGGACGACGA 4450
 D R D R I R R G G F L P I T D D E
 15 GGGCATGGGGATGCAT
 G

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS
 with the endogenous AT domain replaced by the AT domain of module 13 (specific for
 20 methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the
 amino acid sequence shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
 Q L A E A L L T L V R E S T
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
 25 A A V L G H V G G E D I P A T A A
 GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150
 F K D L G I D S L T A V Q L R N
 CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
 A L T E A T G V R L N A T A V F D
 30 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACCTGACCGG 250
 F P T P H V L A G K L G D E L T G
 CACCGCGCGCGCGTCTGCCCCGGACCGCGGCCACGGCCGGTGCACG 300
 T R A P V V P R T A A T A G A H
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGGGGGTC 350
 35 D E P L A I V G M A C R L P G G V
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
 A S P E E L W H L V A S G T D A I
 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTGACGCGATCTACGACC 450
 T E F P T D R G W D V D A I Y D
 40 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
 P D P D A I G K T F V R H G G F L
 ACCGGCGCGACAGGCTTCGACGCGCGTTCCTCGGCATCAGCCCGCGCGA 550
 T G A T G F D A A F F G I S P R E
 GGCCCTCGCGATGGACCCGACGAGCGGGTGTCTCTGGAGACGTCTGTTGG 600
 45 A L A M D P Q Q R V L L E T S W
 AGGCGTTTCGFAAGCGCCGGCATCACCCCGGACTCGACCCGCGGACGCGAC 650
 E A F E S A G I T P D S T R G S D
 ACCGGCGTGTTCGTGCGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700
 T G V F V G A F S Y G Y G T G A D
 50 CACCGACGGCTTCGGCGCGACCGGCTCGCAGACAGTGTGCTCTCCGGCC 750
 T D G F G A T G S Q T S V L S G
 GGCTGTCTGTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTCGACACG 800
 R L S Y F Y G L E G P A V T V D T
 GCGTGTTCGTCTGCTGGTGGCGCTGCACCAGGCCGGGAGTCGCTGCG 850
 55 A C S S S L V A L H Q A G Q S L R
 CTCCGGCGAATGCTCGCTCGCCCTGGTCGGCGGCGTCACGGTGATGGCGT 900
 S G E C S L A L V G G V T V M A
 CTCCCGGCGGCTTCGTGGAGTTCTCCCGGACGCGCGGCTCGCGCCGGAC 950
 S P G G F V E F S R Q R G L A P D
 60 GGCCGGTCGAAGGCGTTTCGGCGCGGGTGGCGACGGCACGAGCTTCGCCGA 1000

G R A K A F G A G A D G T S F A E
GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
G A G V L I V E R L S D A E R N
GTCACACCGTCTGGCGGTGCTCCGTGGTTCGGCGGTCAACCAGGATGGT 1100
5 G H T V L A V V R G S A V N Q D G
GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTGCGCAGGAGCGGGTGAT 1150
A S N G L S A P N G P S Q E R V I
CCGGCAGGCCCCTGGCCAACGCCGGGCTCACCCCGGCGGACGTGGACGCCG 1200
R Q A L A N A G L T P A D V D A
10 TCGAGGCCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGCACAG 1250
V E A H G T G T R L G D P I E A Q
GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTGCTGGG 1300
A V L A T Y G E R A T P L L L G
CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGCTCGCCG 1350
15 S L K S N I G H A Q A A S G V A
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400
G I I K M V Q A L R H G E L P P T
CTGCACGCCGACGAGCCGTCGCCGCACGTCGACTGGACGGCCGGCGCCGCT 1450
L H A D E P S P H V D W T A G A V
20 CGAAGTGTGCTGACGTGCGCCCGGCGCTGGCCCGAGACCGACCGGCCACGGC 1500
E L L T S A R P W P E T D R P R
GTGCCGCCGCTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCACGTCATC 1550
R A A V S S F G V S G T N A H V I
CTGGAGGCCGGACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGA 1600
25 L E A G P V T E T P A A S P S G D
CCTTCCCCTGCTGGTGTGCGGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650
L P L L V S A R S P E A L D E Q
TCCGCCGACTGCGCGCCTACCTGGACACCACCCCGGACGTCGACCGGGTG 1700
I R R L R A Y L D T T P D V D R V
30 GCCGTGGCACAGACGCTGGCCCCGGCGCACACTTCGCCCCACCGCGCCGT 1750
A V A Q T L A R R T H F A H R A V
GCTGCTCGGTGACACCGTCATCACACACCCCGCGGACCGGCCGACG 1800
L L G D T V I T T P P A D R P D
AACTCGTCTTCTGCTACTCCGGCCAGGGCACCCAGCATCCCGCATGGGC 1850
35 E L V F V Y S G Q G T Q H P A M G
GAGCAGCTAGCCGATTCTGTCGTTGGTGTTCGCCGAGCGGATGGCCGAGTG 1900
E Q L A D S S V V F A E R M A E C
TGCGGCGGCGTTCGCGGAGTTCTGTTGACTGGGATCTGTTACGGTTCTGG 1950
A A A L R E F V D W D L F T V L
40 ATGATCCGGCGGTGGTGACCGGTTGATGTGGTCCAGCCCGCTTCTGG 2000
D D P A V V D R V D V V Q P A S W
GCGATGATGGTTTCCCTGGCCGCGGTGTGGCAGGCGCGCGGTGTGCGGCC 2050
A M M V S L A A V W Q A A G V R P
GGATGCGGTGATCGGCCATTTCGAGGGTGAGATCGCCGACGCTTGTGTGG 2100
45 D A V I G H S Q G E I A A A C V
CGGGTGGGTGCTACTACGCGATGCCGCCCGGATCGTGACCTTGGCGAGC 2150
A G A V S L R D A A R I V T L R S
CAGGCGATCGCCCGGGGCTGGCGGGCGGGGCGCGATGGCATCCGTCGC 2200
Q A I A R G L A G R G A M A S V A
50 CCTGCCCGCGCAGGATGTGAGCTGGTTCGACGGGGCCTGGATCGCCGCC 2250
L P A Q D V E L V D G A W I A A
ACAACGGGCCCCGCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTGAC 2300
H N G P A S T V I A G T P E A V D
CATGTCCTCACCGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGGATCAC 2350
55 H V L T A H E A Q G V R V R R I T
CGTCGACTATGCCTCGCACACCCCGCACGTCGAGCTGATCCGCGACGAAC 2400
V D Y A S H T P H V E L I R D E
TACTCGACATCACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGG 2450
L L D I T S D S S S Q T P L V P W
60 CTGTGACCGTGGACGGCACCTGGGTGACAGCCCGCTGGACGGGGAGTA 2500
L S T V D G T W V D S P L D G E Y
CTGGTACCGGAACCTGCGTGAACCGGTGCGTTTCCACCCCGCCGTCAGCC 2550
W Y R N L R E P V G F H P A V S
AGTTGCAGGCCCAGGGCGACACCGTGTTCGTCGAGGTGACGCCAGCCCCG 2600

Q L Q A Q G D T V F V E V S A S P
 GTGTTGTTGTCAGGCGATGGACGACGATGTCGTCACGGTTGCCACGCTGCG 2650
 V L L Q A M D D D V V T V A T L R
 5 TCGTGACGACGGCGACGCCACCCGGATGCTCACC GCCCTGGCACAGGCCT 2700
 R D D G D A T R M L T A L A Q A
 ATGTCCACGGCGTCACCGTCGACTGGCCCGCCATCCTCGGCACCACCACA 2750
 Y V H G V T V D W P A I L G T T T
 ACCCGGGTACTGGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTG 2800
 T R V L D L P T Y A F Q H Q R Y W
 10 GCTCGAGTCGGGACGCCCCGGCCGATCCGACGCGGGCCACCCCGTGCTGG 2850
 L E S A R P A A S D A G H P V L
 GCTCCGGTATCGCCCTCGCCGGGTCGCGGGGCCGGGTGTTACGGGTTCC 2900
 G S G I A L A G S P G R V F T G S
 GTGCCGACCGGTGCGGACCGCGCGGTGTTCTGTCGCCGAGCTGGCGCTGGC 2950
 15 V P T G A D R A V F V A E L A L A
 CGCCGCGGACGCGGTGCTGCGCCACGGTCGAGCGGCTCGACATCGCCT 3000
 A A D A V D C A T V E R L D I A
 CCGTGCCCGGCCGGGCCATGGCCGGACGACCGTACAGACCTGGGTC 3050
 S V P G R P G H G R T T V Q T W V
 20 GACGAGCCGGCGGACGACGGCCGGCGCGGTTCACCGTGCACACCCGCAC 3100
 D E P A D D G R R R F T V H T R T
 CGGCGACGCCCCGTGGACGCTGCACGCCGAGGGGGTGCTGCGCCCCCATG 3150
 G D A P W T L H A E G V L R P H
 GCACGGCCCTGCCGATGCGGCCGACGCCGAGTGGCCCCCACC GGCGCG 3200
 25 G T A L P D A A D A E W P P P G A
 GTGCCCGCGGACGGGCTGCCGGGTGTGTGGCGCCGGGGGGACAGGTCTT 3250
 V P A D G L P G V W R R G D Q V F
 CGCCGAGGCCGAGGTGGACGGACCGGACGGTTTCGTGGTGCACCCCGACC 3300
 A E A E V D G P D G F V V H P D
 30 TGCTGACGCGGTCTTCTCCGCGGTGCGCGACGGAAGCCGCCAGCCGGCC 3350
 L L D A V F S A V G D G S R Q P A
 GGATGGCGCGACCTGACGGTGCACGCGTCGGACGCCACCGTACTGCGCGC 3400
 G W R D L T V H A S D A T V L R A
 CTGCCTCACCCGGCGCACCGACGGAGCCATGGGATTGCGCCGCTTCGACG 3450
 35 C L T R R T D G A M G F A A F D
 GCGCCGGCCTGCCGGTACTCACCGCGGAGGCGGTGACGCTGCGGGAGGTG 3500
 G A G L P V L T A E A V T L R E V
 GCGTCACCGTCCGGCTCCGAGGAGTCGGACGGCCTGCACCGGTTGGAGTG 3550
 A S P S G S E E S D G L H R L E W
 40 GCTCGCGGTGCGCGAGGCGGTCTACGACGGTGACCTGCCCGAGGGACATG 3600
 L A V A E A V Y D G D L P E G H
 TCCTGATCACCGCCGCCACCCCGACGACCCCGAGGACATACCCACCCGC 3650
 V L I T A A H P D D P E D I P T R
 GCCCACACCCGCGCCACCCGCTCCTGACCGCCTGCAACACCACCTCAC 3700
 45 A H T R A T R V L T A L Q H H L T
 CACCACCGACCACACCTCATCGTCCACACCACCACCGACCCCGCCGGCG 3750
 T T D H T L I V H T T T D P A G
 CCACCGTCACCGGCTCACCCGCACCGCCAGAACGAACACCCCCACCGC 3800
 A T V T G L T R T A Q N E H P H R
 50 ATCCGCTCATCGAAACCGACACCCCCACACCCCTCCCTCCCTGGCCCA 3850
 I R L I E T D H P H T P L P L A Q
 ACTCGCCACCTCGACCACCCCCACCTCCGCTCACCACACACCTCC 3900
 L A T L D H P H L R L T H H T L
 ACCACCCCCACCTACCCCTCCACACCACCACCCACCCACCCACCACC 3950
 55 H H P H L T P L H T T T P P T T T
 CCCCTCAACCCCGAACACGCCATCATCATACCGGCGGCTCCGGCACCCCT 4000
 P L N P E H A I I I T G G S G T L
 CGCCGGCATCCTCGCCCGCCACCTGAACACCCCCACCTACCTCCTCT 4050
 A G I L A R H L N H P H T Y L L
 60 CCCGCACCCACCCCGGACGCCACCCCGGACCCACCTCCCTGCGAC 4100
 S R T P P P D A T P G T H L P C D
 GTCGGGACCCCACTCGCCACACCTCACCACATCCCCAACCC 4150
 V G D P H Q L A T T L T H I P Q P
 CCTCACCGCATCTTCCACACCGCCGCCACCTCGACGACGGCATCCTCC 4200

L T A I F H T A A T L D D G I L
 ACGCCCTCACCCCGACCGCCTCACCACCGTCTCCACCCCAAAGCCAAC 4250
 H A L T P D R L T T V L H P K A N
 GCCGCCTGGCACCTGCACCACCTCACCCAAAACCAACCCCTCACCCACTT 4300
 5 A A W H L H H L T Q N Q P L T H F
 CGTCCTCTACTCCAGCGCCGCGCGTCTCGGCAGCCCCGGACAAGGAA 4350
 V L Y S S A A A V L G S P G Q G
 ACTACGCGCGCCCAACGCCTTCCTCGACGCCCTCGCCACCCACCGCCAC 4400
 N Y A A A N A F L D A L A T H R H
 10 ACCCTCGGCCAACCCGCCACCTCCATCGCCTGGGGCATGTGGCACACCAC 4450
 T I G Q P A T S I A W G M W H T T
 CAGCACCTCACCGACAACCTCGACGACGCCGACCGGGACCGCATCCGCC 4500
 S T L T G Q L D D A D R D R I R
 GCGGCGGTTTCTCCGATCACGGACGACGAGGGCATGGGGATGCAT
 15 R G G F L P I T D D E G

Phage KC515 DNA was prepared using the procedure described in Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D. Hopwood *et al.* A phage suspension prepared from 10 plates (100 mm) of confluent plaques of KC515 on *S. lividans* TK24 generally gave about 3 µg of phage DNA. The DNA was ligated to circularize at the cos site, subsequently digested with restriction enzymes *Bam*HI and *Pst*I, and dephosphorylated with SAP.

Each module 8 cassette described above was excised with restriction enzymes *Bgl*II and *Nsi*I and ligated into the compatible *Bam*HI and *Pst*I sites of KC515 phage DNA prepared as described above. The ligation mixture containing KC515 and various cassettes was transfected into protoplasts of *Streptomyces lividans* TK24 using the procedure described in Genetic Manipulation of *Streptomyces*, A Laboratory Manual edited by D. Hopwood *et al.* and overlaid with TK24 spores. After 16-24 hr, the plaques were restreaked on plates overlaid with TK24 spores. Single plaques were picked and resuspended in 200 µL of nutrient broth. Phage DNA was prepared by the boiling method (Hopwood *et al.*, *supra*). The PCR with primers spanning the left and right boundaries of the recombinant phage was used to verify the correct phage had been isolated. In most cases, at least 80% of the plaques contained the expected insert. To confirm the presence of the resistance marker (thiostrepton), a spot test is used, as described in Lomovskaya *et al.* (1997), in which a plate with spots of phage is overlaid with mixture of spores of TK24 and phiC31 TK24 lysogen. After overnight incubation, the plate is overlaid with antibiotic in soft agar. A working stock is made of all phage containing desired constructs.

Streptomyces hygroscopicus ATCC 14891 (see US Patent No. 3,244,592, issued 5 Apr 1966, incorporated herein by reference) mycelia were infected with the recombinant phage by mixing the spores and phage (1×10^8 of each), and incubating on R2YE agar (Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D.

Hopwood *et al.*) at 30°C for 10 days. Recombinant clones were selected and plated on minimal medium containing thiostrepton (50 µg/ml) to select for the thiostrepton resistance-conferring gene. Primary thiostrepton resistant clones were isolated and purified through a second round of single colony isolation, as necessary. To obtain
5 thiostrepton-sensitive revertants that underwent a second recombination event to evict the phage genome, primary recombinants were propagated in liquid media for two to three days in the absence of thiostrepton and then spread on agar medium without thiostrepton to obtain spores. Spores were plated to obtain about 50 colonies per plate, and thiostrepton sensitive colonies were identified by replica plating onto thiostrepton
10 containing agar medium. The PCR was used to determine which of the thiostrepton sensitive colonies reverted to the wild type (reversal of the initial integration event), and which contain the desired AT swap at module 8 in the ATCC 14891-derived cells. The PCR primers used amplified either the KS/AT junction or the AT/DH junction of the wild-type and the desired recombinant strains. Fermentation of the recombinant strains,
15 followed by isolation of the metabolites and analysis by LCMS, and NMR is used to characterize the novel polyketide compounds.

Example 2

Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-506

20 The present invention also provides the 13-desmethoxy derivatives of FK-506 and the novel PKS enzymes that produce them. A variety of *Streptomyces* strains that produce FK-506 are known in the art, including *S. tsukubaensis* No. 9993 (FERM BP-927), described in U.S. Patent No. 5,624,852, incorporated herein by reference; *S. hygroscopicus* subsp. *yakushimaensis* No. 7238, described in U.S. patent No. 4,894,366,
25 incorporated herein by reference; *S. sp.* MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference; and *S. sp.* MA 6548, described in Motamedi *et al.*, 1998, "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506," *Eur. J. Biochem.* 256: 528-534, and Motamedi *et al.*,
1997, "Structural organization of a multifunctional polyketide synthase involved in the
30 biosynthesis of the macrolide immunosuppressant FK-506," *Eur. J. Biochem.* 244: 74-80, each of which is incorporated herein by reference.

The complete sequence of the FK-506 gene cluster from *Streptomyces sp.* MA6548 is known, and the sequences of the corresponding gene clusters from other FK-506-producing organisms is highly homologous thereto. The novel FK-506 recombinant
35 gene clusters of the present invention differ from the naturally occurring gene clusters in

that the AT domain of module 8 of the naturally occurring PKSs is replaced by an AT domain specific for malonyl CoA or methylmalonyl CoA. These AT domain replacements are made at the DNA level, following the methodology described in Example 1.

- 5 The naturally occurring module 8 sequence for the MA6548 strain is shown below, followed by the illustrative hybrid module 8 sequences for the MA6548 strains.

```

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
  M R L Y E A A R R T G S P V V V
  GCGGCCGCECTCGACGACGCGCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
10  A A A L D D A P D V P L L R G L R
  GCGTACGACCGTCCGGCGGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150
  R T T V R R A A V R E R S L A D
  GCTCGCCGTGCTGCCGACGACGAGCGCGCCGACGCCTCCCTCGCGTTTCG 200
  R S P C C P T T S A P T P P S R S
15  TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
  S W N S T A T V L G H L G A E D I
  CCCGGCGACGACGACGTTCAGGAACCTCGGCATCGACTCGCTCACC GCG 300
  P A T T T F K E L G I D S L T A
  TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350
20  V Q L R N A L T T A T G V R L N A
  ACAGCGGTCTTCGACTTTCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400
  T A V F D F P T P R A L A A R L G
  CGACGAGCTGGCCGGTACCCGCGCGCCCGTCCGGCCCCGACCGCGGCCA 450
  D E L A G T R A P V A A R T A A
25  CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
  T A A A H D E P L A I V G M A C R
  CTGCCGGGCGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
  L P G G V A S P Q E L W R L V A S
  CGGCACCGCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
30  G T D A I T E F P A D R G W D V
  ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
  D A L Y D P D P D A I G K T F V R
  CACGGCGGCTTCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
  H G G F L D G A T G F D A A F F G
35  GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
  I S P R E A L A M D P Q R V L
  TGGAGACGTCTGGGAGGCGTTTCGAAAGCGCGGCATCACCCCGGACGCG 800
  L E T S W E A F E S A G I T P D A
  GCGCGGGGACGCGACACCGGCGTGTTCATCGGCGGTTCTCCTACGGGTA 850
40  A R G S D T G V F I G A F S Y G Y
  CGGCACGCGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGCA 900
  G T G A D T N G F G A T G S Q T
  GCGTGCTCTCCGGCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
  S V L S G R L S Y F Y G L E G P S
45  GTCACGGTCGACACCGCCTGCTCGTCGTCACTGGTCGCCCTGCACCAGGC 1000
  V T V D T A C S S S L V A L H Q A
  AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050
  G Q S L R S G E C S L A L V G G
  TCACGGTGATGGCGTCGCCCGGCGATTTCGTCGAGTTCTCCCGGACGCG 1100
50  V T V M A S P G G F V E F S R Q R
  GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
  G L A P D G R A K A F G A G A D G
  TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTCGAGCGGCTCTCCG 1200
  T S F A E G A G A L V V E R L S
55  ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250
  D A E R H G H T V L A L V R G S A
  GCTAACTCCGACGGCGCGTCAACGGTCTGTGCGGCGCCGAACGGCCCTC 1300
  A N S D G A S N G L S A P N G P S
  CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAACTCACCCCGG 1350

```


Q E R V I H Q A L A N A K L T P
 CCGATGTCGACGCGGTGAGGCGCACGGCACCGGACCCGCTCGGGCGAC 1400
 A D V D A V E A H G T G T R L G D
 5 CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
 P I E A Q A L L A T Y G Q D R A T
 GCCCCTGCTGCTCGGCTCGTGAAGTCGAACATCGGGCACGCCAGGCCG 1500
 P L L L G S L K S N I G H A Q A
 CGTCAGGGGTGCGCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
 A S G V A G I I K M V Q A I R H G
 10 GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCCGACGTGCGACTG 1600
 E L P P T L H A D E P S P H V D W
 GACGGCCGGTCCGCTCGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGA 1650
 T A G A V E L L T S A R P W P G
 CCGGTGCGCCGCGCGCGCTGCCGTCTCGTCTCGGCGTGAGCGGCACG 1700
 15 T G R P R R A A V S S F G V S G T
 AACGCCCACATCATCTTGAGGCAGGACCGGTCAAAACGGGACCGGTGCGA 1750
 N A H I I L E A G P V K T G P V E
 GGCAGGAGCGATCGAGGCAGGACCGGTGCAAGTAGGACCGGTGCGAGGCTG 1800
 A G A I E A G P V E V G P V E A
 20 GACCGCTCCCCGCGCGCGCGCGCTCAGCACCGGGCGAAGACCTTCCGCTG 1850
 G P L P A A P P S A P G E D L P L
 CTCGTGTGCGGCGGTTCCCCGGAGGCACTCGACGAGCAGATCGGGCGCCT 1900
 L V S A R S P E A L D E Q I G R L
 GCGCGCCTATCTCGACACCGGCCCGGGCGTGCACCGGGCGGCCGTGGCGC 1950
 25 R A Y L D T G P G V D R A A V A
 AGACACTGGCCCGGCGTACGCACTTCACCCACCGGGCCGTACTGCTCGGG 2000
 Q T L A R R T H F T H R A V L L G
 GACACCGTCATCGGCGCTCCCCCGCGGACCAGGCCGACGAACTCGTCTT 2050
 D T V I G A P P A D Q A D E L V F
 30 CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAACTCG 2100
 V Y S G Q G T Q H P A M G E Q L
 CGGCCGCGTTCCCCGTGTTCCCGATGCCTGGCACGACGCGCTCCGACGG 2150
 A A A F P V F A D A W H D A L R R
 CTCGACGACCCCGACCCGACGACCCACACGGAGCCAGCACACGCTCTT 2200
 35 L D D P D P H D P T R S Q H T L F
 CGCCACACAGGCGGCGTTACCGCCCTCCTGAGGTCTGGGACATCACGC 2250
 A H Q A A F T A L L R S W D I T
 CGCAGCCCGTCATCGGCCACTCGCTCGGCGAGATCACCGCCGCGTACGCC 2300
 P H A V I G H S L G E I T A A Y A
 40 GCCGGGATCCTGTGCTCGACGACGCTGCACCCTGATCACACGCGTGC 2350
 A G I L S L D D A C T L I T T R A
 CCGCCTCATGCACACGCTTCCGCGCGCCGCGCATGGTCACCGTGCTGA 2400
 R L M H T L P P P G A M V T V L
 CCAGCGAGGAGGAGGCCCGTCAGGCGCTGCGGCCGGGCGTGGAGATCGCC 2450
 45 T S E E E A R Q A L R P G V E I A
 GCGGTCTTCGGCCCGCACTCCGTGCTGCTCTCGGGCGACGAGGACGCCGT 2500
 A V F G P H S V V L S G D E D A V
 GCTCGACGTGCACACGCGCTCGGCATCCACCACCGTCTGCCCCGCGCCG 2550
 L D V A Q R L G I H H R L P A P
 50 ACGCGGGCCACTCCGCGCACATGGAACCCGTGGCCGCGGAGCTGCTCGCC 2600
 H A G H S A H M E P V A A E L L A
 ACCACTCGCGAGCTCCGTTACGACCGGCCACACCGCCATCCCGAACGA 2650
 T T R E L R Y D R P H T A I P N D
 CCCCACACCGCCGAGTACTGGGCGGAGCAGGTCCGCAACCCCGTGCTGT 2700
 55 P T T A E Y W A E Q V R N P V L
 TCCACGCCCACACCCAGCGGTACCCCGACCGGTGTTGCTCGAGATCGGC 2750
 F H A H T Q R Y P D A V F V E I G
 CCCGGCCAGGACCTCTCACCCTGGTGCAGGCGATCGCCCTGCAGAACGG 2800
 P G Q D L S P L V D G I A L Q N G
 60 CACGGCGGACGAGGTGCACGCGTGCACACCGCGCTCGCCCGCCTCTTCA 2850
 T A D E V H A L H T A L A R L F
 CACGCGGCGCCACGCTCGACTGGTCCCGCATCCTCGGCGGTGCTTCGCGG 2900
 T R G A T L D W S R I L G G A S R
 CACGACCCTGACGTCCCCTCGTACGCGTTCAGCGGCGTCCCTACTGGAT 2950

H D P D V P S Y A F Q R R P Y W I
CGAGTCGGCTCCCCGGCCACGGCCGACTCGGGCCACCCCGTCCTCGGCA 3000
E S A P P A T A D S G H P V L G
5 CCGGAGTCGCGTTCGCGGGTTCGCGGGCGGGTGTTCACGGTCCCGTG 3050
T G V A V A G S P G R V F T G P V
CCCGCCGGTTCGGACCGCGCGGTGTTCATCGCCGAACGGCGCTCGCCGC 3100
P A G A D R A V F I A E L A L A A
CGCCGACGCCACCGACTGCGCCACGGTTCGAACAGCTCGACGTCACCTCCG 3150
A D A T D C A T V E Q L D V T S
10 TGCCCGGGCGGATCCGCGCGGGCAGGGCCACCGCGCAGACCTGGGTTCGAT 3200
V P G G S A R G R A T A Q T W V D
GAACCCGCGCGGACGGGGCGGCGCGCTTCACCGTCCACACCCGCGTCGG 3250
E P A A D G R R R F T V H T R V G
CGACGCCCGTGGACGCTGCACGCCGAGGGGGTCTCCGCCCCGGCCGCG 3300
15 D A P W T L H A E G V L R P G R
TGCCCGAGCCGAAGCCGTCGACACCGCCTGGCCCCCGCGGGCGCGGTG 3350
V P Q P E A V D T A W P P P G A V
CCCGCGGACGGGCTGCCGCGGGCGTGGCGACGCGCGGACCAGGTCTTCGT 3400
P A D G L P G A W R R A D Q V F V
20 CGAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGC 3450
E A E V D S P D G F V A H P D L
TCGACGCGGTCTTCTCCGCGGTTCGGCGACGGGAGCCGCCAGCCGACCGGA 3500
L D A V F S A V G D G S R Q P T G
25 TGGCGCGACCTCGCGGTGCACGCGTTCGGACGCCACCGTGTCTGCGCGCCTG 3550
W R D L A V H A S D A T V L R A C
CCTCACCCGCGCGACAGTGGTGTCTGGAGCTCGCCGCGCTTCGACGGTG 3600
L T R R D S G V V E L A A F D G
CCGGAATGCCGGTGTCTACCGCGGAGTTCGGTGACGCTGGGCGAGGTTCGCG 3650
A G M P V L T A E S V T L G E V A
30 TCGCGAGGCGGATCCGACGAGTTCGGACGGTCTGCTTCGGCTTGAGTGGTT 3700
S A G G S D E S D G L L R L E W L
GCCGGTGGCGGAGGCCCCACTACGACGGTGGCGACGAGCTGCCCCGAGGGCT 3750
P V A E A H Y D G A D E L P E G
35 ACACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCACCAAC 3800
Y T L I T A T H P D D P D D P T N
CCCCACAACACCCACACGACCCACACACAAACCACACGCGTCCTCAC 3850
P H N T P T R T H T Q T T R V L T
CGCCCTCCAACACCACCTCATCACCAACCAACACCCCTCATCGTCCACA 3900
A L Q H H L I T T N H T L I V H
40 CCACCACCGACCCCCAGGCGCGCGGTCACCGGCCTCACCCGACCGCA 3950
T T T D P P G A A V T G L T R T A
CAAAACGAACCCCCGGCGCATCCACCTCATCGAAACCCACACCCCCA 4000
Q N E H P G R I H L I E T H H P H
45 CACCCCACTCCCCCTCACCCAACTACCAACCCCTCCACCAACCCACCTAC 4050
T P L P L T Q L T T L H Q P H L
GCCTCACCAACAACACCCCTCCACACCCCCACCTACCCCATCACCAAC 4100
R L T N N T L H T P H L T P I T T
CACCACAACACCACAACCAACCCCAACACCCCAACCCCTCAACCCAA 4150
H H N T T T T T P N T P P L N P N
50 CCACGCCATCCTCATACCGGCGGCTCCGGCACCCCTCGCGGCATCCTCG 4200
H A I L I T G G S G T L A G I L
CCCGCCACCTCAACACCCCCACACCTACCTCCTCTCCGACACACCACCA 4250
A R H L N H P H T Y L L S R T P P
55 CCCCCACACACCCGGCACCCACATCCCTGCGACCTACCGACCCAC 4300
P P T T P G T H I P C D L T D P T
CCAAATCACCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCT 4350
Q I T Q H I P Q P L T G I
TCCACACCGCGCCACCCCTCGACGACGCCACCCCTACCAACCTCACCCCC 4400
F H T A A T L D D A T L T N L T P
60 CAACACCTCACCAACCCCTCCAACCCAAAGCCGACGCGCCTGGCACCT 4450
Q H L T T T L Q P K A D A A W H L
CCACCACACACCCAAAACCAACCCCTCACCACTTCGTCTCTACTCCA 4500
H H H T Q N Q P L T H F V L Y S
GCGCCGCGCCACCCCTCGGCAGCCCCGGCCAAGCCAACCTACGCGCGCGCC 4550

S A A A T L G S P G Q A N Y A A A
 AACGCCTTCCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACC 4600
 N A F L D A L A T H R H T Q G Q P
 CGCCACCACCATCGCCTGGGGCATGTGGCACACCACCACTCACCA 4650
 5 A T T I A W G M W H T T T T L T
 GCCAACTCACCGACAGCGACCGCGACCGCATCCGCCGCGGCGGCTTCCTG 4700
 S Q L T D S D R D R I R R G G F L
 CCGATCTCGGACGACGAGGGCATGC
 10 P I S D D E G M

The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
 M R L Y E A A R R T G S P V V V
 15 GCGGCCGCGCTCGACGACGCGCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
 A A A L D D A P D V P L L R G L R
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150
 R T T V R R A A V R E R S L A D
 20 GCTCGCCGTGCTGCCCCGACGACGAGCGCGCCGACGCTCCCTCGCGTTG 200
 R S P C C P T T S A P T P P S R S
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
 S W N S T A T V L G H L G A E D I
 CCGGCGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300
 P A T T T F K E L G I D S L T A
 25 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGCGTACGCCTCAACGCC 350
 V Q L R N A L T T A T G V R L N A
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400
 T A V F D F P T P R A L A A R L G
 CGACGAGCTGGCCGGTACCCGCGCGCCGTCGCGGCCCGGACCGCGGCCA 450
 30 D E L A G T R A P V A A R T A A
 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
 T A A A H D E P L A I V G M A C R
 CTGCCGGGCGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
 L P G C V A S P Q E L W R L V A S
 35 CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
 G T D A I T E F P A D R G W D V
 ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
 D A L Y D P D P D A I G K T F V R
 CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
 40 H G G F L D G A T G F D A A F F G
 GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGAGCAACGGGTGCTCC 750
 I S P R E A L A M D P Q Q R V L
 TGGAGACGTCTGGGAGGCGTTTCAAAGCGCGGGCATCACCCCGGACGCG 800
 L E T S W E A F E S A G I T P D A
 45 GCGCGGGGACGCGACACCGGCGTGTTCATCGGCGGTTCTCCTACGGGTA 850
 A R G S D T G V F I G A F S Y G Y
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGAGACCA 900
 G T G A D T N G F G A T G S Q T
 GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
 50 S V L S G R L S Y F Y G L E G P S
 GTCACGGTCGACACCGCCTGCTCGTCTGCTACTGGTCGCCCTGCACAGGC 1000
 V T V D T A C S S S L V A L H Q A
 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050
 G Q S L R S G E C S L A L V G G
 55 TCACGGTGATGGCGTCGCCCCGGGATTTCGTGAGTTCTCCCGGACGCGC 1100
 V T V M A S P G G F V E F S R Q R
 GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150
 G L A P D G R A K A F G A G A D G
 TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTCGAGCGGCTCTCCG 1200
 60 T S F A E G A G A L V V E R L S
 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250
 D A E R H G H T V L A L V R G S A

GCTAACTCCGACGGCGCGTCGAACGGTCTGTGCGCGCCGAACGGCCCCCTC 1300
 A N S D G A S N G L S A P N G P S
 CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAACTCACCCTCCG 1350
 Q E R V I H Q A L A N A K L T P
 5 CCGATGTCGACGCGGTGAGGCGCACGGCACCCGCGCTCGGCGAC 1400
 A D V D A V E A H G T G T R L G D
 CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
 P I E A Q A L L A T Y G Q D R A T
 10 GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500
 P L L L G S L K S N I G H A Q A
 CGTCAGGGGTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
 A S G V A G I I K M V Q A I R H G
 GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCGCACGTGACTG 1600
 E L P P T L H A D E P S P H V D W
 15 GACGGCCGGTGCCGTCGAGCTCCTGACGTGCGCCCCGGCCGTGGCCGGGGA 1650
 T A G A V E L L T S A R P W P G
 CCGGTGCGCCCTAGGCGGGCAGGCGTGTCTCTTCGGGATCAGTGCCACC 1700
 T G R P R R A G V S S F G I S G T
 AACGCCACGTCATCCTGGAAAGCGCACCCCCCACTCAGCCTGCGGACAA 1750
 20 N A H V I L E S A P P T Q P A D N
 CGCGGTGATCGAGCGGGCACCGGAGTGGGTGCCGTTGGTGATTTCGGCCA 1800
 A V I E R A P E W V P L V I S A
 GGACCCAGTCGGCTTTGACTGAGCACGAGGGCCGGTTGCGTGCGTATCTG 1850
 R T Q S A L T E H E G R L R A Y L
 25 GCGGCGTCGCCCCGGGTGGATATGCGGGCTGTGGCATCGACGCTGGCGAT 1900
 A A S P G V D M R A V A S T L A M
 GACACGGTCCGTGTTTCGAGCACCGTGCCGTGCTGCTGGGAGATGACACCG 1950
 T R S V F E H R A V L L G D D T
 30 TCACCGGCACC3CTGTGTCTGACCCTCGGGCGGTGTTTCGTCTTCCCGGGA 2000
 V T G T A V S D P R A V F V F P G
 CAGGGGTGCGCAGCGTGCTGGCATGGGTGAGGAACTGGCCGCGCGTTCCC 2050
 Q G S Q R A G M G E E L A A A F P
 CGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCCG 2100
 V F A R I H Q Q V W D L L D V P
 35 ATCTGGAGGTGAACGAGACCGGTTACGCCCAGCCGGCCCTGTTTCGCAATG 2150
 D L E V N E T G Y A Q P A L F A M
 CAGGTGGCTCTGTTCGGGCTGCTGGAATCGTGGGGTGTACGACCGGACGC 2200
 Q V A L F G L L E S W G V R P D A
 GGTGATCGGCCATTCCGGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGG 2250
 40 V I G H S V G E L A A A Y V S G
 TGTGGTCTGTGGAGGATGCCTGCACTTTGGTGTGCGGCGGGGCTCGTCTG 2300
 V W S L E D A C T L V S A R A R L
 ATGACGGCTCTGCCCGCGGGTGGGGTGTGGTGTGCTGTCGCGGTCTCGGA 2350
 M Q A L P A G G V M V A V P V S E
 45 GGATGAGGCCCGGGCCGTGCTGGGTGAGGGTGTGGAGATCGCCGCGGTCA 2400
 D E A R A V L G E G V E I A A V
 ACGGCCCGCTCGTEGGTGGTTCTCTCCGGTGATGAGGCCGCGGTGCTGCAG 2450
 N G P S S V V L S G D E A A V L Q
 50 GCCGCGGAGGGGCTGGGGAAGTGGACGCGGCTGGCGACCAGCCACGCGTT 2500
 A G E L G K W T R L A T S H A F
 CCATTCCGCCCCGTATGGAACCCATGCTGGAGGAGTTCCGGGCGGTGCGCG 2550
 H S A R M E P M L E E F R A V A
 AAGGCCTGACCTACCGGACGCGCAGGTCTCCATGGCCGTTGGTGATCAG 2600
 E G L T Y R T P Q V S M A V G D Q
 55 GTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGGACACGGTCCGTT 2650
 V T T A E Y W V R Q V R D T V R F
 CGGCGAGCAGGTGGCCTCGTACGAGGACGCGGTGTTTCGTGAGCTGGGTG 2700
 G E Q V A S Y E D A V F V E L G
 CCGACCGGTCACTGGCCCGCTGGTTCGACGGTGTGCGGATGCTGCACGGC 2750
 60 A D R S L A R L V D G V A M L H G
 GACCACGAAATCCAGGCCGCGATCGGGCCCTGGCCACCTGTATGTCAA 2800
 D H E I Q A A I G A L A H L Y V N
 CGGCGTCACGGTGCAGTGGCCCGCGCTCCTGGGCGATGCTCCGGCAACAC 2850
 G V T V D W P A L L G D A P A T

GGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCAGCGCTACTGGCTC 2900
R V L D L P T Y A F Q H Q R Y W L
GAGTCGGCTCCCCGGGCCACGGCCGACTCGGGCCACCCCGTCTCGGCAC 2950
E S A P P A T A D S G H P V L G T
5 CGGAGTCGCGGTGCGCGGGTGGCGGGCGGGTGTTCACGGGTCCCGTGC 3000
G V A V A G S P G R V F T G P V
CCGCGGTGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCC 3050
P A G A D R A V F I A E L A L A A
10 GCCGACGCCACCGACTGCGCCACGGTTCGAACAGCTCGACGTCACCTCCGT 3100
A D A T D C A T V E Q L D V T S V
GCCCGGGCGGATCCGCGCGGCGAGGGCCACCGCGCAGACCTGGGTGATG 3150
P G C S A R G R A T A Q T W V D
AACCCGCGCGGACGGGCGCGCTTACCGTCCACACCCGCGTCCGGC 3200
E P A A D G R R R F T V H T R V G
15 GACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCCGGCGCGT 3250
D A P W T L H A E G V L R P G R V
GCCCGAGCCCGAAGCCGTGACACCGCCTGGCCCCCGCGGGCGCGGTGC 3300
P Q Q P E A V D T A W P P P G A V
CCGCGGACCGCTGCGGGGGCGTGGCGACGCGCGGACCAGGTCTTCGTC 3350
20 P A D G L P G A W R R A D Q V F V
GAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCT 3400
E A E V D S P D G F V A H P D L L
CGACGCGGTCTTCTCCGCGGTGCGCGACGGGAGCCGCCAGCCGACCGGAT 3450
D A V F S A V G D G S R Q P T G
25 GCGCGACCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCCTGC 3500
W R D L A V H A S D A T V L R A C
CTCACCCGCGCGACAGTGGTGTCTGGAGCTCGCCGCCTTCGACGGTGC 3550
L T R R D S G V V E L A A F D G A
30 CGGAATGCCGGTGTCTACCGCGGAGTCCGGTACGCTGGGCGAGGTGCGGT 3600
G M P V L T A E S V T L G E V A
CGGCAGGCGGATCCGACGAGTCCGACGGTCTGCTTCGGCTTGAGTGGTTG 3650
S A G G S D G L R L E W L
CCGGTGGCGGAGGCCACTACGACGGTGCCGACGAGCTGCCCGAGGGGCTA 3700
P V A E A H Y D G A D E L P E G Y
35 CACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCACCAACC 3750
T L I T A T H P D D P D D P T N
CCCACAACACACACACGACCCACACACAAACCACACGCGTCCCTCACC 3800
P H N T P T R T H T Q T T R V L T
GCCCTCCAACACCACTCATCACCACCAACCAACCTCATCGTCCACAC 3850
40 A L Q H H L I T T N H T L I V H T
CACCACCGACCCCGAGGCGCGCGGTACCGGCTCACCCTGACCCGACCCGAC 3900
T T D P P G A A V T G L T R T A
AAAACGAACACCCCGCGCATCCACCTCATCGAAACCCACCAACCCAC 3950
Q N E H P G R I H L I E T H H P H
45 ACCCACTCCCCCTACCCAACCTCACCACCTCCACCAACCCACCTACG 4000
T P L T Q L T T L H Q P H L R
CCTCACCAACAACACCTCCACACCCCGACCTCACCCTCATCACCACCC 4050
L T N N T L H T P H L T P I T T
ACCACAACACCACACACACACCCCAACACCCACCCCTCAACCCCAAC 4100
50 H H N T T T T T P N T P P L N P N
CACGCCATCCTCATCACCAGGCGGCTCCGGCACCCCTCGCCGGCATCCTCGC 4150
H A I L I T G G S G T L A G I L A
CGGCACCTCAACCACCCCGACCTACCTCCTCTCCCGCACACCCAC 4200
R H L N H P H T Y L L S R T P P
55 CCCCCACCAACCCCGGACCCACATCCCTGCGACCTCACCAGCCCCACC 4250
P P T T P G T H I P C D L T D P T
CAAATCACCACAGCCCTCACCACATACCACAACCCCTCACCAGGATCTT 4300
Q I T Q A L T H I P Q P L T G I F
CCACACCGCGCCACCTCGACGACGCCACCTCACCACCTCACCACCC 4350
60 H T A A T L D D A T L T N L T P
AACACCTCACCACACCCCTCCAACCCAAAGCCGACGCGCGCTGGCACCTC 4400
Q H L T T T L Q P K A D A A W H L
CACCACCAACCCAAACCAACCCCTCACCACCTCGTCTCTACTCCAG 4450
H H H T Q N Q P L T H F V L Y S S

CGCCGCCGCCACCCCTCGGCAGCCCCGGCCAAGCCAACTACGCCGCCGCCA 4500
 A A A T L G S P G Q A N Y A A A
 ACGCCTTCCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCC 4550
 N A F L D A L A T H R H T Q G Q P
 5 GCCACCACCATCGCCTGGGGCATGTGGCACACCACCACCACTCACCAG 4600
 A T T I A W G M W H T T T T L T S
 CCAACTCACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCCTGC 4650
 Q L T D S D R D R I R R G G F L
 CGATCTCGGACGACGAGGGCATGC
 10 P I S D D E G M

The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
 15 M R L Y E A A R R T G S P V V V
 GCGGCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
 A A A L D D A P D V P L L R G L R
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150
 R T T V R R A A V R E R S L A D
 20 GCTCGCCGTGCTGCCCAGCAGAGCGCGCGACGCCTCCCTCGCGTTCG 200
 R S P C P T T S A P T P P S R S
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
 S W N S T A T V L G H L G A E D I
 CCCGGCGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300
 25 P A T T T F K E L G I D S L T A
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350
 V Q L R N A L T T A T G V R L N A
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCTCGCCGCGAGACTCGG 400
 T A V F D F P T P R A L A A R L G
 30 CGACGAGCTGCGCGGTACCCGCGCGCCCGTCCGCGCCCGGACCGCGGCCA 450
 D E L A G T R A P V A A R T A A
 CCGCGGCCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
 T A A A H D E P L A I V G M A C R
 CTGCGGGCGGGGTGCGTCCGACAGGAGCTGTGGCGTCTCGTCGCGTC 550
 35 L P G G V A S P Q E L W R L V A S
 CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
 G T D A I T E F P A D R G W D V
 ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
 D A L Y D P D P D A I G K T F V R
 40 CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCCTCGG 700
 H G F L D G A T G F D A A F F G
 GATCAGCCCCGCGGAGGCCCTGGCCATGGACCCGAGCAACGGGTGCTCC 750
 I S P R E A L A M D P Q Q R V L
 TGGAGACGTCTGGGAGGCGTTCGAAAGCGGGGCATACCCCGGACGCG 800
 45 L E T S W E A F E S A G I T P D A
 GCGCGGGGCGAGCGACACCGGCGTGTTCATCGGCGCGTTCCTACGGGTA 850
 A R G S D T G V F I G A F S Y G Y
 CGGCACGGGTGCGGATACCAACGGGCTTCGGCGCGACAGGGTCGCAGACCA 900
 G T G A D T N G F G A T G S Q T
 50 GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
 S V L S G R L S Y F Y G L E G P S
 GTCACGGTCGACACCGCCTGCTCGTCGTCACCTGGTCGCCCTGCACCAGGC 1000
 V T V D T A C S S S L V A L H Q A
 AGGGAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCGGCGGTG 1050
 55 G Q S L R S G E C S L A L V G G
 TCACGGTGATGGCGTCGCCCCGGCGGATTCGTGAGTTCTCCCGGCAGCGC 1100
 V T V M A S P G G F V E F S R Q R
 GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150
 G L A P D G R A K A F G A G A D G
 60 TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200
 T S F A G A L V V E R L S
 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250

D A E R H G H T V L A L V R G S A
 GCTAACTCCGACGGCGCTCGAACGGTCTGTGCGGCGCCGAACGGCCCCTC 1300
 A N S D G A S N G L S A P N G P S
 5 CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350
 Q E R V I H Q A L A N A K L T P
 CCGATGTGACGCGGTGAGGCGCACGGCACCGGCACCGCCTCGGCGAC 1400
 A D V D A V E A H G T G T R L G D
 CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
 P I E A Q A L L A T Y G Q D R A T
 10 GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500
 P L L L G S L K S N I G H A Q A
 CGTCAGGGGTGCGCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
 A S G V A G I I K M V Q A I R H G
 GAACTGCCGCCGACACTGCACGCGGACGAGCCGTCGCCGCACGTCTGACTG 1600
 15 E L P P T L H A D E P S P H V D W
 GACGGCCGGTGCCGTCGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGGA 1650
 T A G A V E L L T S A R P W P G
 CCGTTCGCCCCTAGGCGGGCGGGCTGTCTCCTTCGGAGTCAGCGGCACC 1700
 T G R P R R A G V S S F G V S G T
 20 AACGCCCCACGTCATCCTGGAGAGCGCACCCCCCGCTCAGCCCCGCGGAGGA 1750
 N A H V I L E S A P P A Q P A E E
 GCGCGAGCCTGTTGAGACGCCGGTGGTGGCCTCGGATGTGCTGCCGCTGG 1800
 A Q P V E T P V V A S D V L P L
 25 TGATATCGGCCAAGACCCAGCCCCGCCCTGACCGAACACGAAGACCGGCTG 1850
 V I S A K T Q P A L T E H E D R L
 CGCGCCTACCTGGCGGCTCGCCCCGGGCGGATATACGGGCTGTGGCATC 1900
 R A Y L A A S P G A D I R A V A S
 GACGCTGGCGGTGACACGGTTCGGTGTTCGAGCACCGCGCCGTACTCCTTG 1950
 T L A V T R S V F E H R A V L L
 30 GAGATGACACCGTCAACCGGCACCGCGGTGACCGACCCAGGATCGTGTTT 2000
 G D D T V T G T A V T D P R I V F
 GTCTTTCCCGGCGAGGGGTGGCAGTGGCTGGGGATGGGCAGTGCAGTGGC 2050
 V F P G Q G W L G M G S A L R
 CGATTCTGTCGGTGGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCGT 2100
 35 D S S V V F A E R M A E C A A A
 TGCGCGAGTTTCGTGGACTGGGATCTGTTACGGTTCCTGGATGATCCGGCG 2150
 L R E F V D W D L F T V L D D P A
 GTGGTGGACCGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGGT 2200
 V V D V V Q P A S W A M M V
 40 TTCCCTCGCCGCGGTGTGGCAGGCGGCGGTGTGCGGCCGATGCGGTGA 2250
 S L A A V W Q A A G V R P D A V
 TCGGCCATTTCGAGGTGAGATCGCCGACGCTTGTGTGGCGGGTGGCGTG 2300
 I G H S Q G E I A A A C V A G A V
 TCACTACGCGATGCCGCCCGGATCGTGACCTTGCGCAGCCAGGCGATCGC 2350
 45 S L R D A R I V T L R S Q A I A
 CCGGGGCTGGCGGGCGGGCGGATGGCATCCGTGCGCCCTGCGCGCGC 2400
 R G L A G R G A M A S V A L P A
 AGGATGTGAGCTGGTTCGACGGGGCCTGGATCGCCGCCACAACGGGGCCC 2450
 Q D V E L V D G A W I A A H N G P
 50 GCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTGACCATGTCTCTCAC 2500
 A S T V I A G T P E A V D H V L T
 CGCTCATGAGGCACAAGGGGTGCGGGTGGCGGATCACCGTCGACTATG 2550
 A H E A Q G V R V R I T V D Y
 CCTCGCACACCCCGCACGTGAGCTGATCCGCGACGAACACTACTCGACATC 2600
 55 A S H T P H V E L I R D E L L D I
 ACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTGACCGT 2650
 T S D S S S Q T P L V P W L S T V
 GGACGGCACCTGGGTGACAGCCCGCTGGACGGGGAGTACTGGTACCGGA 2700
 D G T W V D S P L D G E Y W Y R
 60 ACCTGCGTGAACCGGTGCGTTTCCACCCCGCGTCAGCCAGTTGACAGGCC 2750
 N L R E P V G F H P A V S Q L Q A
 CAGGGCGACACCGTGTTCGTGAGGTGAGCGCCAGCCCGGTGTTGTTGCA 2800
 Q G D T V F V E V S A S P V L L Q
 GGCGATGGACGACGATGTCGTACGGTTGCCACGCTGCGTCTGACGACG 2850

A M D D D V V T V A T L R R D D
GCGACGCCACCCGGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGGC 2900
G D A T R M L T A L A Q A Y V H G
5 GTCACCGTCGACTGGCCCGCCATCCTCGGCACCACCACAACCCGGGTACT 2950
V T V D W P A I L G T T T T R V L
GGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGG 3000
D L P T Y A F Q H Q R Y W L E S
CTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCCTCGGCACCCGGAGTC 3050
A P P A T A D S G H P V L G T G V
10 GCCGTCGCCGGGTGCGCCGGCCGGGTGTTACGGGTCCCGTGCCCGCCGG 3100
A V A G S P G R V F T G P V P A G
TGCGGACCGCGGGTGTTCATCGCCGAACCTGGCGCTCGCCGCCGCCGACG 3150
A D R A V F I A E L A L A A A D
CCACCGACTGCGCCACGGTGAACAGCTCGACGTCACCTCCGTGCCCGGC 3200
15 A T D C A T V E Q L D V T S V P G
GGATCCGCCCCGCGCAGGGCCACCGCGCAGACCTGGGTGATGAACCCGC 3250
G S A R G R A T A Q T W V D E P A
CGCCGACGGCGCGCCGCTTACCGTCCACACCCGCGTCGGCGACGCCC 3300
A D G R R R F T V H T R V G D A
20 CGTGGACGCTGCACGCCGAGGGGTTCTCCGCCCGGCCGCGTGGCCCGAG 3350
P W T L H A E G V L R P G R V P Q
CCCGAAGCCGTCGACACCGCCTGGCCCCCGCCGGGCGCGGTGCCCGCGGA 3400
P E A V D T A W P P P G A V P A D
CGGGCTGCCCGGGGCGTGGCGACGCGCGGACCAGGTCTTCGTGAAGCCG 3450
25 G L P G A W R R A D Q V F V E A
AAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCTCGACGCG 3500
E V D S P D G F V A H P D L L D A
GTCTTCTCCGCGGTGCGCGACGGGAGCCGCCAGCCGACCGGATGGCGCGA 3550
V F S A V G D G S R Q P T G W R D
30 CCTCGCGGTGCACGCGTCGGACGCCACCGTGCTGCGCGCTGCCCTACCC 3600
L A V H A S D A T V L R A C L T
GCCGCGACAGTGGTGTCTCGTGGAGCTCGCCGCTTCGACGGTGCCGGAATG 3650
R R D S G V V E L A A F D G A G M
CCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTGCGGTGCGGCAGG 3700
35 P V L T A E S V T L G E V A S A G
CGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTGG 3750
G S D E S D G L L R L E W L P V
CGGAGGCCCCACTACGACGGTGCCGACGAGTGGCCGAGGGCTACACCCCTC 3800
A E A H Y D G A D E L P E G Y T L
40 ATCACCGCCACACACCCCGACGACCCCGACGACCCCAACACCCCAACAA 3850
I T A T H P D D P D D P T N P H N
CACACCCACACGACCCACACACAAACCACACGCGTCTCACCGCCCTCC 3900
T P T R T H T Q T T R V L T A L
AACACCACTCATCACCAACCAACCACTCATCGTCCACACCAACCACC 3950
45 Q H H L I T T N H T L I V H T T T
GACCCCCCAGGCGCGCGCGTCAACCGCCTCAACCGCACCGCACAAAACGA 4000
D P P G A A V T G L T R T A Q N E
ACACCCCGCGCATCCACCTCATCGAAACCCACCAACCCCAACACCCAC 4050
H P G R I H L I E T H H P H T P
50 TCCCCCTCACCAACTCACCAACCTCCACCAACCCCACTACGCCTCACC 4100
L P L T Q L T T L H Q P H L R L T
AACAACACCTCCACACCCCACTCACCCTCATCACCAACCAACCAAA 4150
N N T L H T P H L T P I T T H H N
CACCACCAACCAACCCCAACACCCCAACCCCTCAACCCCAACCAACGCCA 4200
55 T T T T T P N T P P L N P N H A
TCCTCATCACCGGCGGCTCCGGCACCCCTCGCCGCGATCCTCGCCCGCCAC 4250
I L I T G G S G T L A G I L A R H
CTCAACCAACCCCACTACCTCCTCTCCCGCACCAACCAACCCCAAC 4300
L N H P T Y L L S R T P P P P T
60 CACACCCGGCACCCACATCCCTGCGACCTACCGACCCCAACCAATCA 4350
T P G T H I P C D L T D P T Q I
CCCAAGCCCTCACCAACATACCAACCCCTACCGGCATCTTCCACACC 4400
T Q A L T H I P Q P L T G I F H T
GCCGCCACCCCTCGACGACGCCACCCCTACCAACCTCACCCCAACACCT 4450

A A T L D D A T L T N L T P Q H L
 CACCACCACCTCCAACCCAAAGCCGACGCCGCTGGCACCTCCACCACC 4500
 T T T L Q P K A D A A W H L H H
 ACACCCAAAACCAACCCCTCACCCTTCGTCCTCTACTCCAGCGCCGCC 4550
 5 H T Q N Q P L T H F V L Y S S A A
 GCCACCTCTCGGCAGCCCCGGCCAAGCCAACTACGCCGCCGCCAACGCCTT 4600
 A T L G S P G Q A N Y A A A N A F
 CCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCCGCCACCA 4600
 L D A L A T H R H T Q G Q P A T
 10 CCATCGCCTGGGGCATGTGGCACACCACCACCACTCACCAGCCAACTC 4700
 T I A W G M W H T T T T L T S Q L
 ACCGACAGCGACCGACCGCATCCGCCGCGGGCTTCCTGCCGATCTC 4750
 T D S D R D R I R R G G F L P I S
 GGACGACGAGGGCATGC
 15 D D E G M

The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
 20 M R L Y E A A R R T G S P V V V
 GCGGCGCGCTCGACGACGCGCCGACGTGCCGCTGCTGCGCGGGCTGCG 100
 A A A L D D A P D V P L R G L R
 GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150
 R T T V R R A A V R E R S L A D
 25 GCTCGCCGTGCTGCCCCGACGACGAGCGCGCCGACGCCTCCCTCGCGTTCTG 200
 R S P C C P T T S A P T P P S R S
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
 S W N S T A T V L G H L G A E D I
 CCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACC GCGG 300
 30 P A T T T F K E L G I D S L T A
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGCGTACGCCTCAACGCC 350
 V Q L P N A L T T A T G V R L N A
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCTCGCCGCGAGACTCGG 400
 T A V F F P T P R A L A A R L G
 35 CGACGAGCTGCGCGGTACCCGCGCGCCCGTCCGCGGCCCGGACCGCGGCCA 450
 D E L A G T R A P V A A R T A A
 CCGCGCGCGCGCACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
 T A A A H D E P L A I V G M A C R
 CTGCCGCGCGGGGTCGCGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
 40 L P G G V A S P Q E L W R L V A S
 CGGACCCGACGCCATCACGGAGTTCCCGCGGACCGCGGCTGGGACGTGG 600
 G T D A I T E F P A D R G W D V
 ACGCGCTCTACGACCCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
 D A L Y D P D P D A I G K T F V R
 45 CACGGCGGCTTCCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
 H G G F L D G A T G F D A A F F G
 GATCAGCCCGCGCGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
 I S P R E A L A M D P Q Q R V L
 TGGAGACGTCTCGGGAGGCGTTCCGAAAAGCGCGGGCATCACCCCGGACGCG 800
 50 L E T S W E A F E S A G I T P D A
 GCGCGGGGCGACACCGGCGTGTTCATCGGCGGTTCTCCTACGGGTA 850
 A R G S D T G V F I G A F S Y G Y
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTTCGACACCA 900
 G T G A D T N G F G A T G S Q T
 55 GCGTGCTCTCCGGCCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
 S V L S G R L S Y F Y G L E G P S
 GTCACGGTCGACACCGCCTGCTCGTCGTCCTGGTTCGCCCTGCACCAGGC 1000
 V T V D T A C S S S L V A L H Q A
 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050
 60 G Q S L R S G E C S L A L V G G
 TCACGGTGATGGCGTCGCCCGGCGATTTCGTCGAGTTCTCCCGGACGCGC 1100
 V T V M A S P G G F V E F S R Q R

GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
G L A P D G R A K A F G A G A D G
TACGAGCTTCGCGGAGGGCGCCGGTGCCTGGTGGTTCGAGCGGCTCTCCG 1200
T S F A E G A G A L V V E R L S
5 ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250
D A E R H G H T V L A L V R G S A
GCTAACTCCGACGGCGGTGCAACGGTCTGTGCGGCGCGAACGGCCCCCTC 1300
A N S D G A S N G L S A P N G P S
CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACGCGAAACTCACCCCCG 1350
10 Q E R V I H Q A L A N A K L T P
CCGATGTGACGCGGTTCGAGGCGCACGGCACCGGCACCGCCTCGGCGAC 1400
A D V D A V E A H G T G T R L G D
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
P I E A Q A L L A T Y G Q D R A T
15 GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500
P L L L G S L K S N I G H A Q A
CGTCAGGGGTGCGCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
A S G V A G I I K M V Q A I R H G
GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCGCACGTGCGACTG 1600
20 E I P P T L H A D E P S P H V D W
GACGGCCGGTGCCGTGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGGA 1650
T A G A V E L L T S A R P W P G
CCGGTTCGCGCGCGCGCGCTGCCGTCTCGTCTGCGGCGTGAGCGGCACG 1700
T G R A R A A V S S F G V S G T
25 AACGCCCACATCATCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCTGA 1750
N A H I I L E A G P V K T G P V E
GGCAGGAGCGATCGAGGCAGGACCGGTGCAAGTAGGACCGGTGAGGCTG 1800
A G A I E A G P V E V G P V E A
GACCGCTCCCCGCGGCGCGCGCGTCAACCGGGCGAAGACCTTCCGCTG 1850
30 G P L P A A P P S A P G E D L P L
CTCGTGTGCGGCGGTTCGCCGAGGCACTGACGAGCAGATCGGGCGCCT 1900
L V S A R S P E A L D E Q I G R L
GCGCGCCTATCTGACACCGGCGCGGCGTCAACCGGGCGGCGCTGGCGC 1950
R A Y L D T G P G V D R A A V A
35 AGACACTGGCCCGCGGTACGCACTTACCCACCGGGCCGTACTGCTCGGG 2000
Q T L A R R T H F T H R A V L L G
GACACCGTATCGGCGCTCCCCCGCGGACCGGCGGACGAACTCGTCTT 2050
D T V I G A P P A D Q A D E L V F
CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100
40 V Y S G Q G T Q H P A M G E Q L
CCGCCGCGTTCGCCGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTG 2150
A A A F P V F A R I H Q Q V W D L
CTCGATGTGCCGATCTGGAGGTGAACGAGACCGGTTACGCCAGCCGGC 2200
L D V P E V N E T G Y A Q P A
45 CCTGTTTCGAATGCAGGTGGCTCTGTTTCGGGCTGCTGGAATCGTGGGGTG 2250
L F A M Q V A L F G L L E S W G
TACGACCGGACGCGGTGATCGGCCATTCCGTGGGTGAGCTTGGCGCTGCG 2300
V R P D A V I G H S V G E L A A A
TATGTGTCCGGGGTGTGGTTCGTTGGAGGATGCCTGCACTTTGGTGTGCGC 2350
50 Y V S G V W S L E D A C T L V S A
GCGGGCTCGTCTGATGCAGGCTCTGCCCGGGTGGGGTGTGCTGCTG 2400
R A R L M Q A L P A G G V M V A
TCCCGGTCTCGGAGGATGAGGCCCGGGCGGTGCTGGGTGAGGGTGTGGAG 2450
V P V S E D E A R A V L G E G V E
55 ATCGCCGCGGTCAACGGCCCGTCTGTCGGTGGTTCTCTCCGGTGTGAGGC 2500
I A A V N G P S S V V L S G D E A
CGCGTGTGTCAGGCGCGGAGGGCTGGGGAAGTGGACGCGGCTGGCGA 2550
A V L Q A A E G G L G K W T R L A
CCAGCCACGCGTTCCATTCCGCCCGTATGGAACCCATGCTGGAGGAGTTC 2600
60 T S H A F H S A R M E P M L E E F
CGGGCGGTGCGCGAAGGCCTGACCTACCGGACGCGCAGGTCTCCATGGC 2650
R A V A E G L T Y R T P Q V S M A
CGTTGGTGTGATCAGGTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGG 2700
V G D Q V T T A E Y W V R Q V R

ACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGACGCCGTGTTC 2750
D T V R F G E Q V A S Y E D A V F
GTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCTGGTTCGACGGTGTTCGC 2800
V E L G A D R S L A R L V D G V A
5 GATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGCGCCCTGGCCC 2850
M L H G D H E I Q A A I G A L A
ACCTGTATGTCAACGGCGTCAACGGTCACTGGCCCGCGCTCCTGGGCGAT 2900
H L Y V N G V T V D W P A L L G D
10 GCTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCA 2950
A P A T R V L D L P T Y A F Q H Q
GCGCTACTGGCTCGAGTCGGCTCCCCCGGCCACGGCCGACTCGGGCCACC 3000
R Y W L E S A P P A T A D S G H
CCGTCTCGGCACCGGAGTCGCCGTCGCCGGGTGCCGGGGCCGGGTGTTC 3050
P V L G T G V A V A G S P G R V F
15 ACGGGTCCCGTGCCCCCGGTGCGGACCGCGCGGTGTTTCATCGCCGAAC 3100
T G P V P A G A D R A V F I A E L
GGCGCTCGCCGCGCCGACGCCACCGACTGCGCCACGGTCAACAGCTCG 3150
A L A A D A A T D C A T V E Q L
ACGTCACTCCGTGCCCGCGGATCCGCCCGCGGCAGGGCCACCGCGCAG 3200
20 D V T S V P G G S A R G R A T A Q
ACCTGGGTGCGATGAACCCGCGCCGACGGGCGGCGCGCTTACCGTCCA 3250
T W V D E P A A D G R R R F T V H
CACCCGCGTTCGGCGACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCC 3300
T R V G D A P W T L H A E G V L
25 GCCCCGGCGCGTGCCCCAGCCCCGAAGCGTTCGACACCGCCTGGCCCCCG 3350
R P G R V P Q P E A V D T A W P P
CCGGGCGCGGTGCCCGCGGACGGGCTGCCCGGGCGTGGCGACGCGCGGA 3400
P G A V P A D G L P G A W R R A D
CCAGGTCTTCGTGCAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCAC 3450
30 Q V F V E A E V D S P D G F V A
ACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTGCGGCGACGGGAGCCGC 3500
H P D L D A V F S A V G D G S R
CAGCCGACCGGATGGCGCGACCTCGCGGTGCACGCGTCGGACGCCACCGT 3550
Q P T G W R D L A V H A S D A T V
35 GCTGCGCGCCTTCCTCACCCGCGCGACAGTGGTGTCTGGAGCTCGCCG 3600
L R A C L T R R D S G V V E L A
CCTTCGACGGTGCCGGAATGCCGGTGTCAACCGCGGAGTCGGTGACGCTG 3650
A F D G A G M P V L T A E S V T L
GGCGAGGTGCGCTCGGCAGGCGGATCCGACGAGTCGGACGGTCTGCTTCG 3700
40 G E V A S A G G S D E S D G L L R
GCTTGAGTGGTTGCCGGTGGCGGAGGCCACTACGACGGTGCCGACGAGC 3750
L E W L P V A E A H Y D G A D E
TGCCCCGAGGGCTACACCCCTCATCACCGCCACACACCCCGACGACCCCGAC 3800
L P E G Y T L I T A T H P D D P D
45 GACCCACCAACCCCAACAACACCCACCGACCCACACACAAACCAC 3850
D P T N P H N T P T R T H T Q T T
ACGCGTCTTCACCGCCCTCCAACACCACCTCATCACCAACCAACCAACCC 3900
R V L T A L Q H H L I T T N H T
TCATCGTCCACACCACCGACCCCCAGGCGCGCGCTCACCGGCCTC 3950
50 L I V H T T T D P P G A A V T G L
ACCCGACCGCACAAAACGAACACCCCGCGCATCCACCTCATCGAAAC 4000
T R T A Q N E H P G R I H L I E T
CCACCACCCCAACCCCACTCCCCCTCACCAACTCACCACTCCACC 4050
H H P H T P L P L T Q L T T L H
55 AACCCACCTACGCCTACCAACAACACCTCCACACCCCCACCTCACC 4100
Q P H L R L T N N T L H T P H L T
CCCATCACCAACCAACAACCAACCAACCAACCAACCAACCAACCAACCAAC 4150
P I T T H H N T T T T T P N T P P
CCTCAACCCCAACCAACCACTCATCACCGGCGGCTCCGGCACCTCG 4200
60 L N P N H A I L I T G G S G T L
CCGGCATCTCGCCCGCCACCTCAACCAACCCCAACCACTACCTCCTCTCC 4250
A G I L A R H L N H P H T Y L L S
CGCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 4300
R T P P P P T T P G T H I P C D L

	CACCGACCCCAACCCAAATCACCCAAGCCCTCACCCACATACCACAACCCC	4350
	T D P T Q I T Q A L T H I P Q P	
	TCACGGCATCTTCCACACCGCGGCCACCCTCGACGACGCCACCCTCACC	4400
	L T G I F H T A A T L D D A T L T	
5	AACCTCACCCCCAACACCTCACCACCACCCTCCAACCCAAAGCCGACGC	4450
	N L T P Q H L T T T L Q P K A D A	
	CGCTGGCACCTCCACCACCACCCAAAACCAACCCCTCACCCACTTCG	4500
	A W H L H H H T Q N Q P L T H F	
10	TCCTCTACTCCAGCGCGCGGCCACCCTCGGACGCCCCGGCCAAGCCAAC	4550
	V L Y S S A A A T L G S P G Q A N	
	TACGCCGCGGCCAACGCCTTCTCGACGCCCTCGCCACCCACCGCCACAC	4600
	Y A A A N A F L D A L A T H R H T	
	CCAAGGACAACCCGCCACCACCATCGCCTGGGGCATGTGGCACACCACCA	4650
	Q G Q P A T T I A W G M W H T T	
15	CCACACTCACCAGCCAACTACCCGACAGCGACCGCGACCGCATCCGCCGC	4700
	T T L T S Q L T D S D R D R I R R	
	GGCGGCTTCTGCGCATCTCGGACGACGAGGGCATGC	
	G G F L P I S D D E G M	

20 The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
M R L Y E A A R R T G S P V V V
25 GCGGCCGCGCTCGACGACGCGCCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
A A A L D D A P D V P L L R G L R
GCGTACGACCGTCCGGCGTGCCGCCGTCCGGGAACGCTCTCTCGCCGACC 150
R T T V R A A V R E R S L A D
GCTCGCCGTGCTGCCCCGACGACGCGCGGACGCTCCCTCGCGTTTCG 200
R S P C C P T T S A P T P P S R S
30 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
S W N S T A T V L G H L G A E D I
CCCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACC GCG 300
P A T T T F K E L G I D S L T A
TCCAGTGCCTCAACGCGCTGACCACCGGACCGGTACGCTCAACGCC 350
V Q L R N A L T T A T G V R L N A
ACAGCGGTCTTCGACTTTCCGACGCGCGCGCTCGCCGCGAGACTCGG 400
T A V F D F P T P R A L A A R L G
CGACGAGCTGGCCGGTACCCGCGCGCCCGTCCGCGCCCGGACCGCGGCCA 450
D E L A G T R A P V A A R T A A
40 CCGGCGCGCGCAGCAGAACCGCTGGCGATCGTGGGCTAGGCTGCCGT 500
T A A A H D E P L A I V G M A C R
CTGCCGGGCGGGGTGCGTTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
L P G G V A S P Q E L W R L V A S
CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
G T D A I T E F P A D R G W D V
ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
D A L Y D P D P D A I G K T F V R
CACGGCGGCTTCCCTCGACGGTTCGACCGGCTTCGACGCGGCTTCTTCGG 700
H G G F L D G A T G F D A A F F G
50 GATCAGCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
I S P R E A L A M D P Q Q R V L
TGGAGACGTCCTGGGAGGCGTTTCGAAAGCGGGGCATCACCCCGGACGCG 800
L E T S W E A A F E S A G G I T P D A
GCGCGGGGCGAGCAGCACCGGCGTGTTCATCGGCGCGTTCTCTACGGGTA 850
55 A R G S D T G V F I G A F S Y G Y
CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGACCA 900
G T G A D T N G F G A T G S Q T
GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
S V L S G R L S Y F Y G L E G P S
60 GTCACGGTCGACACCGCTGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1000
V T V D T A C S S S L V A L H Q A
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050

G Q S L R S G E C S L A L V G G
TCACGGTGATGGCGTCGCGCGGCGGATTCTCGTCGAGTTCTCCCGGCAGCGC 1100
V T V M A S P G G F V E F S R Q R
5 GGGCTCGCGCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
G L A F D G R A K A F G A G A D G
TACGAGCTTCGCCGAGGGCGCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200
T S F A E G A G A L V V E R L S
ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250
D A E R H G H T V L A L V R G S A
10 GCTAACTCCGACGGCGGTGCAACGGTCTGTGCGGCGCCGAACGGCCCTC 1300
A N S D G A S N G L S A P N G P S
CCAGGAACGCGTCATCCACAGGCCCTCGCGAACGCGAACTCACCCCG 1350
Q E R V I H Q A L A N A K L T P
CCGATGTGACGCGGTTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400
15 A D V D A V E A H G T G T R L G D
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
P I E A Q A L L A T Y G Q D R A T
GCCCCGTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500
P L L L G S L K S N I G H A Q A
20 CGTCAGGGGTGCGCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
A S G V A G I I K M V Q A I R H G
GAACTGCCGCCGACACTGCACGCGGACGAGCCGTGCGCGCACGTGACTG 1600
E L P P T L H A D E P S P H V D W
GACGGCCGGTGCCGTGAGCTCCTGACGTGCGCCCGGCCGTGGCCGGGA 1650
25 T A G A V E L L T S A R P W P G
CCGGTCGCCCCGCGCGCTGCCGTCTCGTCGTTTCGGCGTGAGCGGCACG 1700
T G R P R R A A V S S F G V S G T
AACGCCCCACATCATCCTTGAGGCAGGACCGGTCAAACGGGACCGGTGCA 1750
N A H I I L E A G P V K T G P V E
30 GGCAGGAGCGATCGAGGCAGGACCGGTGCAAGTAGGACCGGTGAGGCTG 1800
A G A I E A G P V E V G P V E A
GACCGCTCCCCGCGCGCGCGGTGACACCGGGCGAAGACCTTCCGCTG 1850
G P L P A P A P S A P G E D L P L
CTCGTGTGCGCGGTTCGCCGAGGCACTCGACGAGCAGATCGGGCGCCT 1900
35 L V S A R S P E A L D E Q I G R L
GCGCGCTATCTCGACACCGGCCGCGGTGACCGGGCGGCGCTGGCGC 1950
R A Y L D T G P G V D R A A V A
AGACACTGGCCCGGCGTACGCACTTCACCCACCGGCCGTACTGCTCGGG 2000
Q T L A R R T H F T H R A V L L G
40 GACACCGTCATCGGCGCTCCCCCGGACCGGCGGACGAACTCGTCTT 2050
D T V I G A P P A D Q A D E L V F
CGTC'ACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100
V Y S G Q G T Q H P A M G E Q L
CCGATTCGTGCGGTGGTGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGGC 2150
45 A D S S V V F A E R M A E C A A A
TTGCGCGAGTTCGTGGACTGGGATCTGTTACGGTTCGTGATGATCCGGC 2200
L R E F V D W D L F T V L D D P A
GGTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCTGGGCGATGATGG 2250
V V D R V D V V Q P A S W A M M
50 TTTCCCTGGCCGCGGTGTGGCAGGCGGCGGTGTGCGGCGGATGCGGTG 2300
V S L A A V W Q A A G V R P D A V
ATCGGCCATTTCGAGGGTGAGATCGCCGAGCTTGTGTGGCGGGTGGCGT 2350
I G H S Q G E I A A A C V A G A V
GTCACTACGCGATGCCGCCGCGATCGTGACCTTGGCGAGCCAGGCGATCG 2400
55 S L R D A A R I V T L R S Q A I
CCCGGGGCTTGGCGGGCGGGCGCGATGGCATCCGTGCCCCTGCCCGCG 2450
A R G L A G R G A M A S V A L P A
CAGGATGTGAGCTGGTCGACGGGGCCTGGATCGCCGCCCAACGGGCC 2500
Q D V E L V D G A W I A A H N G P
60 CGCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTGACCATGTCTCTCA 2550
A S T V I A G T P E A V D H V L
CCGCTCATGAGGCACAAGGGGTGCGGGTGGCGGCGATCACCGTCGACTAT 2600
T A H E A Q G V R V R R I T V D Y
GCCTCGCACACCCCGCACGTGAGCTGATCCGCGACGAACTACTCGACAT 2650

A S H T P H V E L I R D E L L D I
 CACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTGACCG 2700
 T S D S S S Q T P L V P W L S T
 TGGACGGCACCTGGGTGACAGCCCGCTGGACGGGGAGTACTGGTACCGG 2750
 5 V D G T W V D S P L D G E Y W Y R
 AACCTGCGTGAACCGGTGCGTTTCCACCCCGCCGTCAGCCAGTTGCAGGC 2800
 N L R E P V G F H P A V S Q L Q A
 CCAGGGCGACACCGTGTTCGTGAGGTGAGCGCCAGCCCGGTGTTGTTGC 2850
 Q G D T V F V E V S A S P V L L
 10 AGGCGATGGACGACGATGTCGTACGGTTGCCACGCTGCGTGTGACGAC 2900
 Q A M D D V D V T V A T L R R D D
 GGCGACGCCACCCGGATGCTCACC GCCCTGGCACAGGCCTATGTCCACGG 2950
 G D A T R M L T A L A Q A Y V H G
 CGTCACCGTGTGACTGGCCCGCCATCTCGGCACCACCACAACCCGGGTAC 3000
 15 V T V D W P A I L G T T T T R V
 TGGACCTTCCGACCTACGCCTTCCAACACCAGCGGTACTGGCTCGAGTCG 3050
 L D L P L T Y A F Q H Q R Y W L E S
 GCTCCCCCGGCACGGCCGACTCGGGCCACCCCGTCTCGGCACCGGAGT 3100
 A P P A T A D S G H P V L G T G V
 20 CGCCGTGCGCGGGTGC CGCGGGCGGGTGTTCACGGGTCCCGTGCCCGCCG 3150
 A V A G S P G R V F T G P V P A
 GTGCGGACCGCGCGGTGTTTCATCGCCGAACTGGCGCTCGCCGCGCCGAC 3200
 G A D R A V F I A E L A L A A A D
 GCCACCGACTGCGCCACGGTGAACAGCTCGACGTACCTCCGTGCCCCG 3250
 25 A T D C A T V E Q L D V T S V P G
 CGGATCCGCCCGCGGCAGGGCCACCGCGCAGACCTGGGTGATGAACCCG 3300
 G S A R G R A T A Q T W V D E P
 CCGCCGACGGGCGGGCGCGCTTACCGTCCACACCCGCGTGGCGACGCC 3350
 A A D G R R R F T V H T R V G D A
 30 CCGTGGACGCTGCACGCCGAGGGGGTCTCCGCCCGCGCGGTGCCCCA 3400
 P W T L H A E G V L R P G R V P Q
 GCCCGAAGCCGTGACACCCGCTGGCCCCCGCGGGCGCGGTGCCCGCGG 3450
 P E A V D T A W P P P G A V P A
 ACGGGCTGCCCCGGGGCGTGCGACGCGCGGACCAGGTCTTCGTGAAGCC 3500
 35 D G L P G A W R R A D Q V F V E A
 GAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCGACCTGCTCGACGC 3550
 E V D S P D G F V A H P D L L D A
 GGTCTTCTCCGCGGTGCGCGACGGGAGCCGCGACCGGATGGCGCG 3600
 V F S A V G D G S R Q P T G W R
 40 ACCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCTGCCTCACC 3650
 D L A V H A S D A T V L R A C L T
 CGCCGCGACAGTGGTGTGCTGGAGCTCGCCGCTTCGACGGTGCCGGAAT 3700
 R R D S G V V E L A A F D G A G M
 GCCCGTGTCTACCGCGGAGTCGGTGACGCTGGGCGAGGTGCGGTGCGCAG 3750
 45 P V L T A E S V T L G E V A S A
 GCGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTGCCGGTG 3800
 G G S D E S D G L L R L E W L P V
 GCGGAGGCCCCACTACGACGGTGCCGACGAGCTGCCCGAGGGCTACACCCT 3850
 A E A H Y D G A D E L P E G Y T L
 50 CATACCGCCACACACCCCGACGACCCCGACGCCACCAACCCCA 3900
 I T A T H P D D P D D P T N P H
 ACACACCCACACGCACCCACACACAAACCACACGCTCCTCACC GCCCTC 3950
 N T P T R T H T Q T T R V L T A L
 CAACACCACCTCATCACCACCAACCACACCCCTCATCGTCCACACCACCAC 4000
 55 Q H H L I T T N H T L I V H T T T
 CGACCCCCCAGGCGCGCGGTACCGGCCTACCCGACCCGACAAAACG 4050
 D P P G A V T G L T R T A Q N
 AACACCCCGCGCATCCACCTCATCGAAACCCACACCCCAACCCCA 4100
 E H P G R I H L I E T H H P H T P
 60 CTCCCCCTACCCAACTCACCACCCCTCCACCAACCCCACTACGCCTCAC 4150
 L P L T Q L T T L H Q P H L R L T
 CAACAACACCCCTCCACACCCCCACCTCACCCTCATCACCACCCACCACA 4200
 N N T L H T P H L T P I T T H H
 ACACCACCAACCAACCCCAACACCCCAACCCCTCAACCCCAACCAACGCC 4250

```

N T T T T T P N T P P L N P N H A
ATCCTCATCACCGGCGGCTCCGGCACCTCGCCGGCATCCTCGCCCGCCA 4300
I L I T G G S G T L A G I L A R H
CCTCAACCACCCACACCTACCTCCTCTCCCGCACACCACCACCCCA 4350
5 L N H P H T Y L L S R T P P P P
CCACACCGGCGACCCACATCCCCTGCGACCTCACCAGCCCCACCCAAATC 4400
T T P G T H I P C D L T D P T Q I
ACCCAAGCCCTCACCACATACCACAACCCCTCACCGGCATCTTCCACAC 4450
T Q A L T H I P Q P L T G I F H T
10 CGCCGCCACCTCGACGACGCCACCCTCACCAACCTCACCCCCAACACC 4500
A A T L D D A T L T N L T P Q H
TCACCACCACCTCCAACCCAAAGCCGACGCCGCTGGCACCTCCACCAC 4550
L T T T L Q P K A D A A W H L H H
CACACCCAAAACCAACCCCTCACCACCTTCGTCCTCTACTCCAGCGCCG 4600
15 H T Q N Q P L T H F V L Y S S A A
CGCCACCTCGGCAGCCCCGGCCAAGCCAACCTACGCCGCCGCCAACGCCT 4650
A T L G S P G Q A N Y A A A N A
TCCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCCGCCACC 4700
F L D A L A T H R H T Q G Q P A T
20 ACCATCGCTGGGGCATGTGGCACACCACCACTCACCAGCCAACT 4750
T I A W G M W H T T T T L T S Q L
CACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCTGCCGATCT 4800
T D S D R D R I R R G G F L P I
CGGACGACGAGGGCATGC
25 S D D E G M

```

Example 3

Recombinant PKS Genes for 13-desmethoxy FK-506 and FK-520

The present invention provides a variety of recombinant PKS genes in addition to those described in Examples 1 and 2 for producing 13-desmethoxy FK-506 and FK-520 compounds. This Example provides the construction protocols for recombinant FK-520 and FK-506 (from *Streptomyces* sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference) PKS genes in which the module 8 AT coding sequences have been replaced by either the *rapAT3* (the AT domain from module 3 of the rapamycin PKS), *rapAT12*, *eryAT1* (the AT domain from module 1 of the erythromycin (DEBS) PKS), or *eryAT2* coding sequences. Each of these constructs provides a PKS that produces the 13-desmethoxy-13-methyl derivative, except for the *rapAT12* replacement, which provides the 13-desmethoxy derivative, i.e., it has a hydrogen where the other derivatives have methyl.

Figure 7 shows the process used to generate the AT replacement constructs. First, a fragment of ~4.5 kb containing module 8 coding sequences from the FK-520 cluster of ATCC 14891 was cloned using the convenient restriction sites *SacI* and *SphI* (Step A in Figure 7). The choice of restriction sites used to clone a 4.0 - 4.5 kb fragment comprising module 8 coding sequences from other FK-520 or FK-506 clusters can be different depending on the DNA sequence, but the overall scheme is identical. The unique *SacI* and *SphI* restriction sites at the ends of the FK-520 module 8 fragment were then changed to unique *Bgl* II and *Nsi* I sites by ligation to synthetic linkers (described in

the preceding Examples, see Step B of Figure 7). Fragments containing sequences 5' and 3' of the AT8 sequences were then amplified using primers, described above, that introduced either an *AvrII* site or an *NheI* site at two different KS/AT boundaries and an *XhoI* site at the AT/DH boundary (Step C of Figure 7). Heterologous AT domains from the rapamycin and erythromycin gene clusters were amplified using primers, as described above, that introduced the same sites as just described (Step D of Figure 7). The fragments were ligated to give hybrid modules with in-frame fusions at the KS/AT and AT/DH boundaries (Step E of Figure 7). Finally, these hybrid modules were ligated into the *BamHI* and *PstI* sites of the KC515 vector. The resulting recombinant phage were used to transform the FK-506 and FK-520 producer strains to yield the desired recombinant cells, as described in the preceding Examples.

The following table shows the location and sequences surrounding the engineered site of each of the heterologous AT domains employed. The FK-506 hybrid construct was used as a control for the FK-520 recombinant cells produced, and a similar FK-520 hybrid construct was used as a control for the FK-506 recombinant cells.

Heterologous AT	Enzyme	Location of Engineered Site
FK-506 AT8 (hydroxymalonyl)	<i>AvrII</i>	GGCCGT <u>ccgcgc</u> CGTGCGGCGGTCTCGTCGTTTC G R P R R A A V S S F
	<i>NheI</i>	ACCCAGCATCCCGCGATGGGTGAGCG <u>gctcgc</u> C T Q H P A M G E R L A
	<i>XhoI</i>	TACGCCTTCCAGCGGCGGCCCTACTGG <u>atcgag</u> Y A F Q R R P Y W I E
rapamycin AT3 (methylmalonyl)	<i>AvrII</i>	GACCGG <u>ccccgc</u> CGGGCGGGCGTGTCTCGTCCTTC D R P R R A G V S S F
	<i>NheI</i>	TGGCAGTGGCTGGGGATGGGCAGTGC <u>cctgcg</u> G W Q W L G M G S A L R
	<i>XhoI</i>	TACGCCTTCCAACACCAGCGGTACTGG <u>gtcgag</u> Y A F Q H Q R Y W V E
rapamycin AT12 (malonyl)	<i>AvrII</i>	GGCCGAG <u>gcgcgc</u> CGGGCAGGCGTGTCTCGTCCTTC G R A R R A G V S S F
	<i>NheI</i>	TCCGAGCGTGTCTGGCATGGGTGAGGA <u>actggc</u> C S Q R A G M G E E L A
	<i>XhoI</i>	TACGCCTTCCAGCACCAGCGCTACTGG <u>gtcgag</u> Y A F Q H Q R Y W L E
DEBS AT1 (methylmalonyl)	<i>AvrII</i>	GCGCGA <u>accgcgc</u> CGGGCGGGGCTCTCGTCGTTTC A R P R R A G V S S F
	<i>NheI</i>	TGGCAGTGGGCGGGCATGGCCGTCGA <u>cctgct</u> C W Q W A G M A V D L L
	<i>XhoI</i>	TACCCGTTCCAGCGCGAGCGCGTCTGG <u>gtcgaa</u> Y P F Q R E R V W L E
DEBS AT2 (methylmalonyl)	<i>AvrII</i>	GACGGG <u>gtgcgc</u> CGGGCAGGTGTGTCTGGCGTTTC D G V R R A G V S A F GCCCAGTGGGAAGGCATGGCGCGGA <u>attatt</u> G

	<i>NheI</i>	A Q W E G M A R E L L
		TATCCTTTCCAGGGCAAGCGGTTCTGG <u>ctgctg</u>
	<i>XhoI</i>	Y P F Q G K R F W L L

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-520 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

5 CCGGCGCCGTCGAACTGCTGACGTCGGCCCCGGCCGTGGCCCCGAGACCGACCGGccacggc
 A G A V E L L T S A R P W P E T D R P R
 GTGCCGCCGTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCACGTCATCCTGGAGGCCG
 R A A V S S F G V S G T N A H V I L E A
 GACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGACCTTCCCCTGCTGGTGTGCG
 10 G P V T E T P A A S P S G D L P L L V S
 CACGCTCACCGGAAGCGCTCGACGAGCAGATCCGCCGACTGCGCGCCTACCTGGACACCA
 A R S P E A L D E Q I R R L R A Y L D T
 CCCCCGAGTCGACCGGGTGGCCGTGGCACAGACGCTGGCCCCGGCGCACACACTTCGCCC
 T P D V D R V A V A Q T L A R R T H F A
 ACCGCGCCGTGCTGCTCGGTGACACCGTCATCACCACACCCCCCGCGGACCGGCCCGACG
 15 H R A V L L G D T V I T T P P A D R P D
 AACTCGTCTTCGTCTACTCCGGCCAGGGCACCCAGCATCCCGCGATGGGCGAGCAgctcg
 E L V F V Y S G Q G T Q H P A M G E Q L
 CGCGCGCCATCCCGTGTTCGCCGACGCCTGGCATGAAGCGCTCCGCCGCCTTGACAACC
 20 A A A P P V F A D A W H E A L R R L D N

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-520 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

25 TCCTCGGGGCTGGGTACGGCACGACGCGGATGTGCCCGCGTACGCGTTCCAACGGCGGC
 I L G A G S R H D A D V P A Y A F Q R R
 ACTACTGGatcgagTCGGCACGCCCCGGCCGATCCGACGCGGGCCACCCCGTGTGGGCT
 H Y W I E S A R P A A S D A G H P V L G

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-506 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

30 TCGGCCAGGCCGTGGCCGCGGACCGGCCGTccgcccCGTGCGCGGCTCTCGTCTCGGG
 S A R P W P R T G R P R R A A V S S F G
 GTGAGCGGCACCAACGCCACATCATCCTGGAGGCCGACCCGACCAGGAGGAGCCGTCG
 35 V S G T N A H I I L E A G P D Q E E P S
 GCAGAACC GGCCGTGACCTCCCGTCTCGTGTGCGGCACGGTCCCGGAGGCACTGGAC
 A E P A G D L P L L V S A R S P E A L D
 GAGCAGATCGGGCGCCTGCGCGACTATCTCGACGCCGCCCCCGCGTGGACCTGGCGGCC
 E Q I G R L R D Y L D A A P G V D L A A
 40 GTGGCGCGGACACTGGCCACGCGTACGCACTTCTCCACCGCGCCGTACTGCTCGGTGAC
 V A R T L A T R T H F S H R A V L L G D
 ACCGTCATCACCGCTCCCCCGTGGAACAGCCGGGCGAGCTCGTCTTCGTCTACTCGGGA
 T V I T A P P V E Q P G E L V F V Y S G
 CAGGGCACCCAGCATCCCGCGATGGGTGAGCGgctcgCGCAGCCTTCCCCGTGTTCGCC
 45 Q G T Q H P A M G E R L A A A F P V F A
 GACCCGACGTACCCGCCTACGCCTTCCAGCGGCGGCCCTACTGGATCGAGTCCGCGCCG
 D P D V P A Y A F Q R R P Y W I E S A P

The sequences shown below provide the location of the AT/DH boundary chosen in the FK-506 module 8 coding sequences. The region where an *XhoI* site was engineered is indicated by lower case and underlining.

GACCCGACGTACCCGCCTACGCCTTCCAGCGGCGGCCCTACTGGatcgagTCCGCGCCG
 D P D V P A Y A F Q R R P Y W I E S A P

Example 4Replacement of Methoxyl with Hydrogen or Methyl at C-15 of FK-506 and FK-520

The methods and reagents of the present invention also provide novel FK-506
 5 and FK-520 derivatives in which the methoxy group at C-15 is replaced by a hydrogen or
 methyl. These derivatives are produced in recombinant host cells of the invention that
 express recombinant PKS enzymes the produce the derivatives. These recombinant PKS
 enzymes are prepared in accordance with the methodology of Examples 1 and 2, with the
 exception that AT domain of module 7, instead of module 8, is replaced. Moreover, the
 10 present invention provides recombinant PKS enzymes in which the AT domains of both
 modules 7 and 8 have been changed. The table below summarizes the various
 compounds provided by the present invention.

	Compound	C-13	C-15	Derivative Provided
15	FK-506	hydrogen	hydrogen	13, 15-didesmethoxy-FK-506
	FK-506	hydrogen	methoxy	13-desmethoxy-FK-506
	FK-506	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-506
	FK-506	methoxy	hydrogen	15-desmethoxy-FK-506
	FK-506	methoxy	methoxy	Original Compound -- FK-506
20	FK-506	methoxy	methyl	15-desmethoxy-15-methyl-FK-506
	FK-506	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-506
	FK-506	methyl	methoxy	13-desmethoxy-13-methyl-FK-506
	FK-506	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-506
	FK-520	hydrogen	hydrogen	13, 15-didesmethoxy FK-520
25	FK-520	hydrogen	methoxy	13-desmethoxy FK-520
	FK-520	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-520
	FK-520	methoxy	hydrogen	15-desmethoxy-FK-520
	FK-520	methoxy	methoxy	Original Compound -- FK-520
	FK-520	methoxy	methyl	15-desmethoxy-15-methyl-FK-520
30	FK-520	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-520
	FK-520	methyl	methoxy	13-desmethoxy-13-methyl-FK-520
	FK-520	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-520

Example 5

35 Replacement of Methoxyl with Ethyl at C-13 and/or C-15 of FK-506 and FK-520

The present invention also provides novel FK-506 and FK-520 derivative compounds in which the methoxy groups at either or both the C-13 and C-15 positions are instead ethyl groups. These compounds are produced by novel PKS enzymes of the invention in which the AT domains of modules 8 and/or 7 are converted to ethylmalonyl specific AT domains by modification of the PKS gene that encodes the module.

Ethylmalonyl specific AT domain coding sequences can be obtained from, for example, the FK-520 PKS genes, the niddamycin PKS genes, and the tylosin PKS genes. The novel PKS genes of the invention include not only those in which either or both of the AT domains of modules 7 and 8 have been converted to ethylmalonyl specific AT domains but also those in which one of the modules is converted to an ethylmalonyl specific AT domain and the other is converted to a malonyl specific or a methylmalonyl specific AT domain.

Example 6

Neurotrophic Compounds

The compounds described in Examples 1 - 4, inclusive have immunosuppressant activity and can be employed as immunosuppressants in a manner and in formulations similar to those employed for FK-506. The compounds of the invention are generally effective for the prevention of organ rejection in patients receiving organ transplants and in particular can be used for immunosuppression following orthotopic liver transplantation. These compounds also have pharmacokinetic properties and metabolism that are more advantageous for certain applications relative to those of FK-506 or FK-520. These compounds are also neurotrophic; however, for use as neurotrophins, it is desirable to modify the compounds to diminish or abolish their immunosuppressant activity. This can be readily accomplished by hydroxylating the compounds at the C-18 position using established chemical methodology or novel FK-520 PKS genes provided by the present invention.

Thus, in one aspect, the present invention provides a method for stimulating nerve growth that comprises administering a therapeutically effective dose of 18-hydroxy-FK-520. In another embodiment, the compound administered is a C-18,20-dihydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18-hydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18,20-dihydroxy-FK-520 derivative. In other embodiments, the compounds are the corresponding analogs of FK-506. The 18-hydroxy compounds of the invention

can be prepared chemically, as described in U.S. Patent No. 5,189,042, incorporated herein by reference, or by fermentation of a recombinant host cell provided by the present invention that expresses a recombinant PKS in which the module 5 DH domain has been deleted or rendered non-functional.

5 The chemical methodology is as follows. A compound of the invention (~200 mg) is dissolved in 3 mL of dry methylene chloride and added to 45 μ L of 2,6-lutidine, and the mixture stirred at room temperature. After 10 minutes, tert-butyldimethylsilyl trifluoromethanesulfonate (64 μ L) is added by syringe. After 15 minutes, the reaction mixture is diluted with ethyl acetate, washed with saturated bicarbonate, washed with
10 brine, and the organic phase dried over magnesium sulfate. Removal of solvent *in vacuo* and flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) gives the protected compound, which is dissolved in 95% ethanol (2.2 mL) and to which is added 53 μ L of pyridine, followed by selenium dioxide (58 mg). The flask is fitted with a water condenser and heated to 70°C on a mantle. After 20 hours, the mixture is
15 cooled to room temperature, filtered through diatomaceous earth, and the filtrate poured into a saturated sodium bicarbonate solution. This is extracted with ethyl acetate, and the organic phase is washed with brine and dried over magnesium sulfate. The solution is concentrated and purified by flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) to give the protected 18-hydroxy compound. This compound is
20 dissolved in acetonitrile and treated with aqueous HF to remove the protecting groups. After dilution with ethyl acetate, the mixture is washed with saturated bicarbonate and brine, dried over magnesium sulfate, filtered, and evaporated to yield the 18-hydroxy compound. Thus, the present invention provides the C-18-hydroxyl derivatives of the compounds described in Examples 1 - 4.

25 Those of skill in the art will recognize that other suitable chemical procedures can be used to prepare the novel 18-hydroxy compounds of the invention. See, e.g., Kawai *et al.*, Jan. 1993, Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues, *FEBS Letters* 316(2): 107-113, incorporated herein by reference. These methods can be used to prepare both the C18-[*S*]-OH and C18-[*R*]-OH enantiomers, with
30 the *R* enantiomer showing a somewhat lower IC₅₀, which may be preferred in some applications. See Kawai *et al.*, *supra*. Another preferred protocol is described in Umbreit and Sharpless, 1977, *JACS* 99(16): 1526-28, although it may be preferable to use 30 equivalents each of SeO₂ and t-BuOOH rather than the 0.02 and 3-4 equivalents, respectively, described in that reference.

All scientific and patent publications referenced herein are hereby incorporated by reference. The invention having now been described by way of written description and example, those of skill in the art will recognize that the invention can be practiced in a variety of embodiments, that the foregoing description and example is for purposes of
5 illustration and not limitation of the following claims.

Claims

1. An isolated nucleic acid that encodes a CoA ligase, a non-ribosomal peptide synthetase, or a domain of an extender module of a polyketide synthase enzyme that synthesizes FK-520.
- 5
2. The isolated nucleic acid of claim 1 that encodes an extender module, said module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.
- 10
3. The isolated nucleic acid of claim 1 that encodes an open reading frame, said open reading frame comprising coding sequences for two or more extender modules, each extender module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.
- 15
4. The isolated nucleic acid of claim 1 that encodes a gene cluster, said gene cluster comprising two or more open reading frames, each of said open reading frames comprising coding sequences for two or more extender modules, each of said extender modules comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.
- 20
5. The isolated nucleic acid of claim 2, wherein at least one of said domains is a domain of a module of a non-FK-520 polyketide synthase.
- 25
6. The isolated nucleic acid of claim 1, wherein said nucleic acid is a recombinant vector capable of replication in or integration into the chromosome of a host cell.
- 30
7. The isolated nucleic acid of claim 6 that is selected from the group consisting of cosmid pKOS034-120, cosmid pKOS034-124, cosmid pKOS065-M27, and cosmid pKOS065-M21.
- 35
8. The isolated nucleic acid of claim 5, wherein said non-FK-520 polyketide synthase is rapamycin polyketide synthase, FK-506 polyketide synthase, or erythromycin polyketide synthase.

9. A method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector of claim 6, and culturing said host cell under conditions such that said polyketide synthase is produced and catalyzes synthesis of said polyketide.

5

10. The method of claim 9, wherein said host cell is a *Streptomyces* host cell.

11. The method of claim 9, wherein said polyketide is selected from the group consisting of FK-520, 13-desmethoxy-FK-520, and 13-desmethoxy-FK-506.

10

12. A recombinant host cell that expresses a recombinant polyketide synthase selected from the group consisting of: (i) an FK-520 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-520 polyketide synthase; (ii) an FK-506 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-506 polyketide synthase; (iii) an FK-520 polyketide synthase in which at least one DH domain has been deleted; (iv) an FK-506 polyketide synthase in which at least one DH domain has been deleted.

15

13. The recombinant host cell of claim 12 that expresses an FK-520 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

20

14. The recombinant host cell of claim 12 that expresses an FK-506 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

25

15. The recombinant host cell of claim 13, wherein a DH domain of module 5 or module 6 has been deleted.

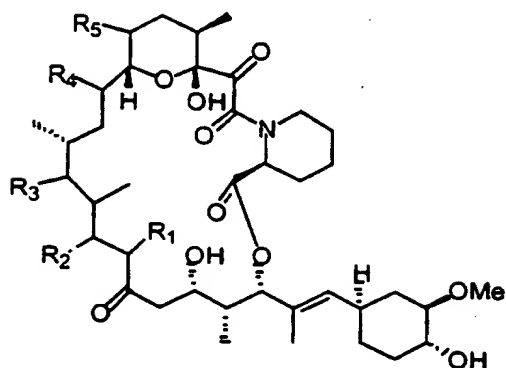
30

16. The recombinant host cell of claim 14, wherein a DH domain of module 5 or module 6 has been deleted.

17. A recombinant host cell that comprises recombinant genes coding for enzymes sufficient for synthesis of ethylmalonyl CoA or 2-hydroxymalonyl CoA.

35

18. A polyketide having the structure



- 5 wherein, R_1 is hydrogen, methyl, ethyl, or allyl; R_2 is hydrogen or hydroxyl, provided that when R_2 is hydrogen, there is a double bond between C-20 and C-19; R_3 is hydrogen or hydroxyl; R_4 is methoxyl, hydrogen, methyl, or ethyl; and R_5 is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506.

10

19. The polyketide of claim 18 that is 13-desmethoxy-FK-506.

20. The polyketide of claim 18 that is 13-desmethoxy-18-hydroxy-FK-520.

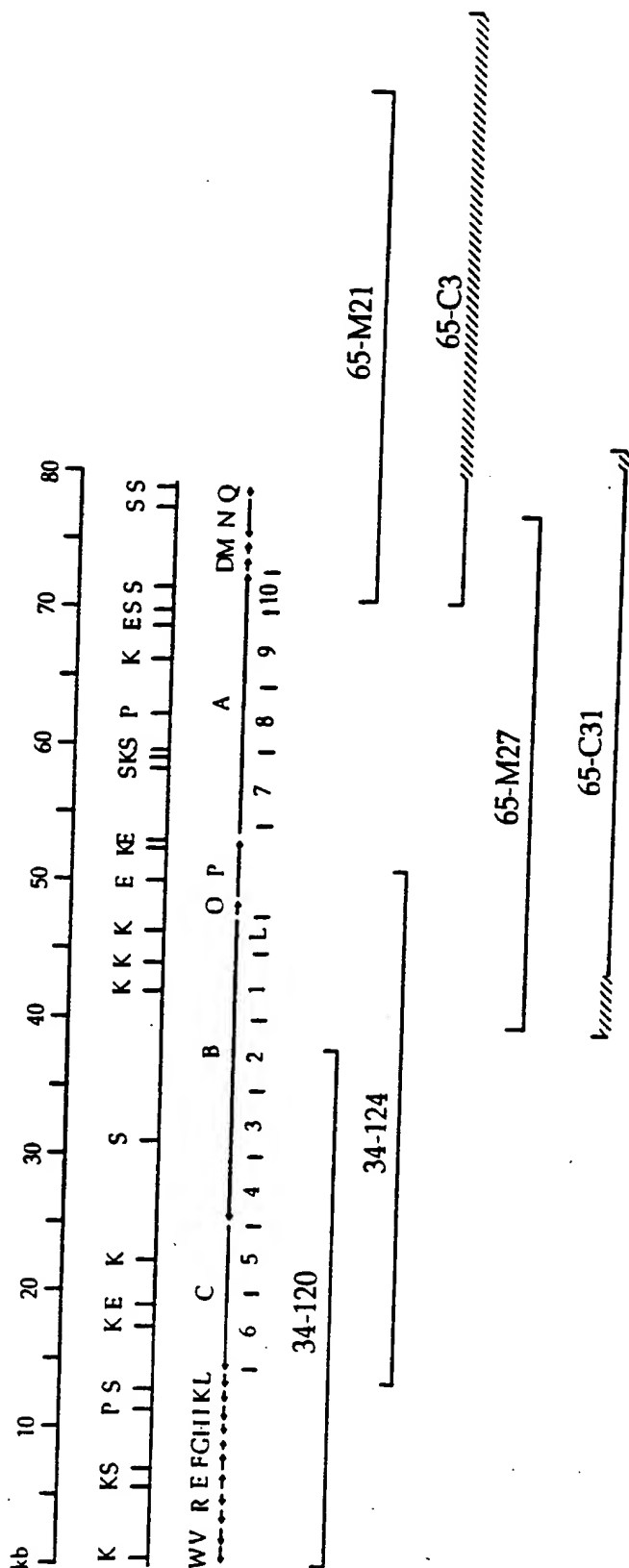


Figure 1

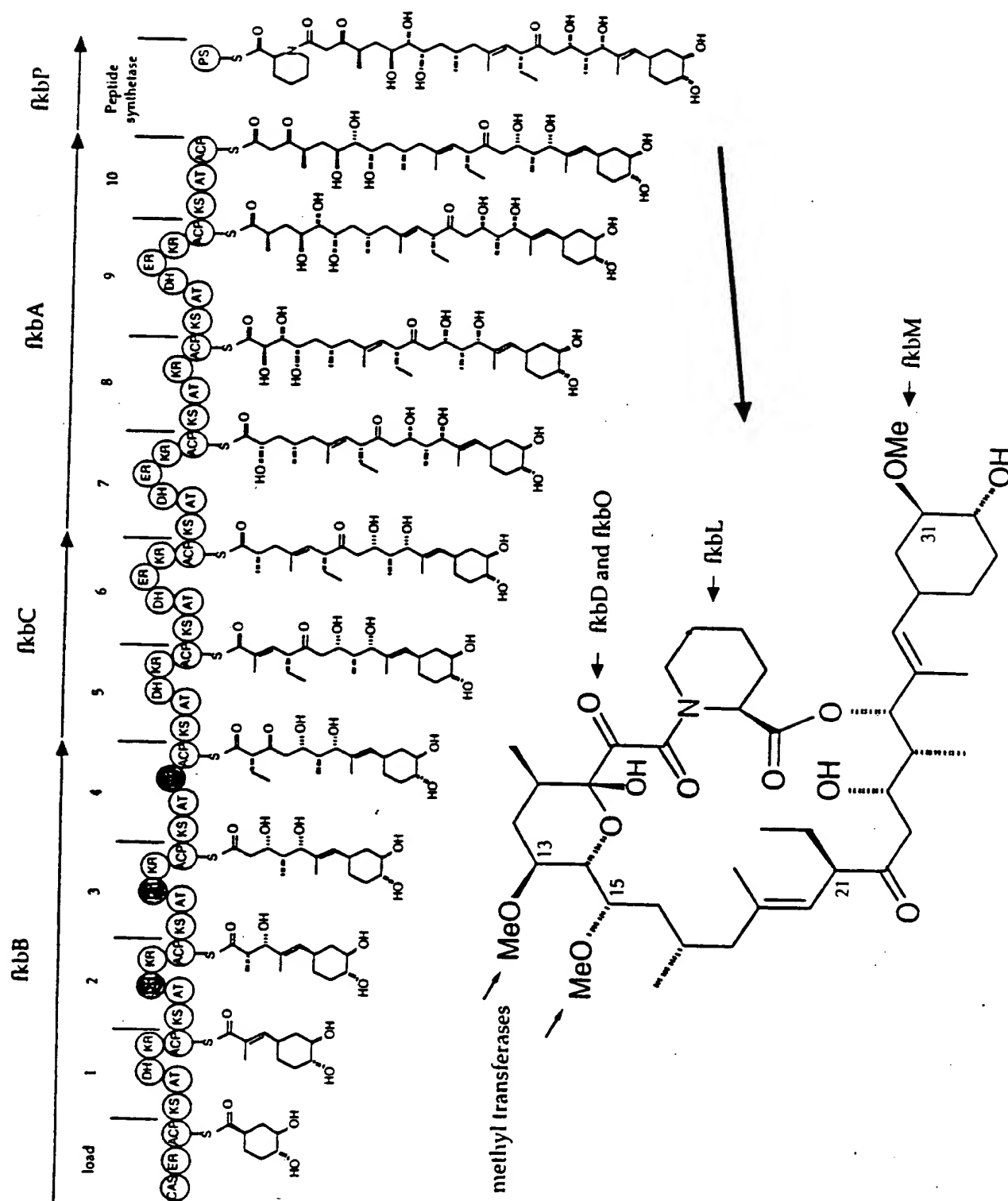


Figure 2

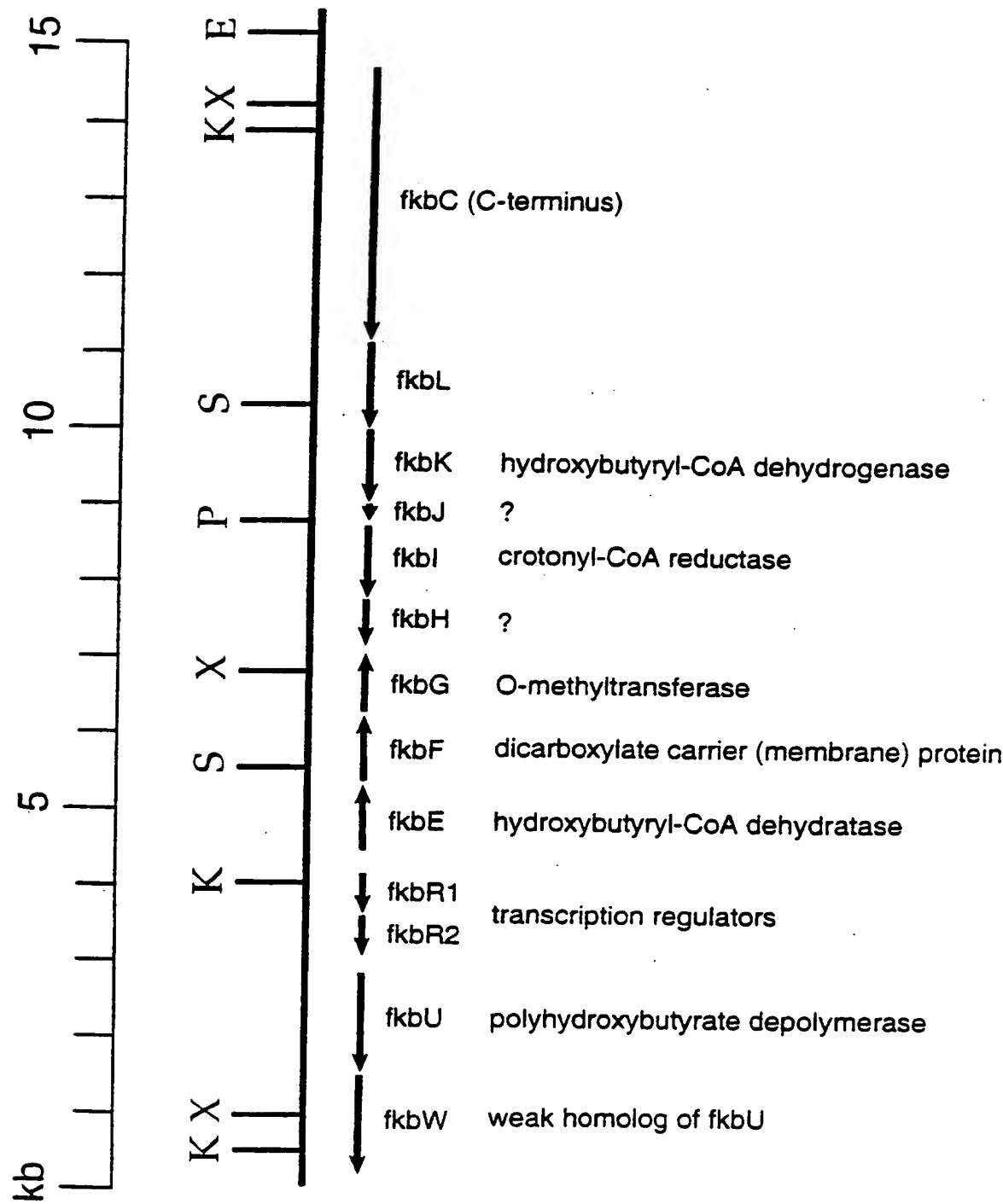


Figure 3

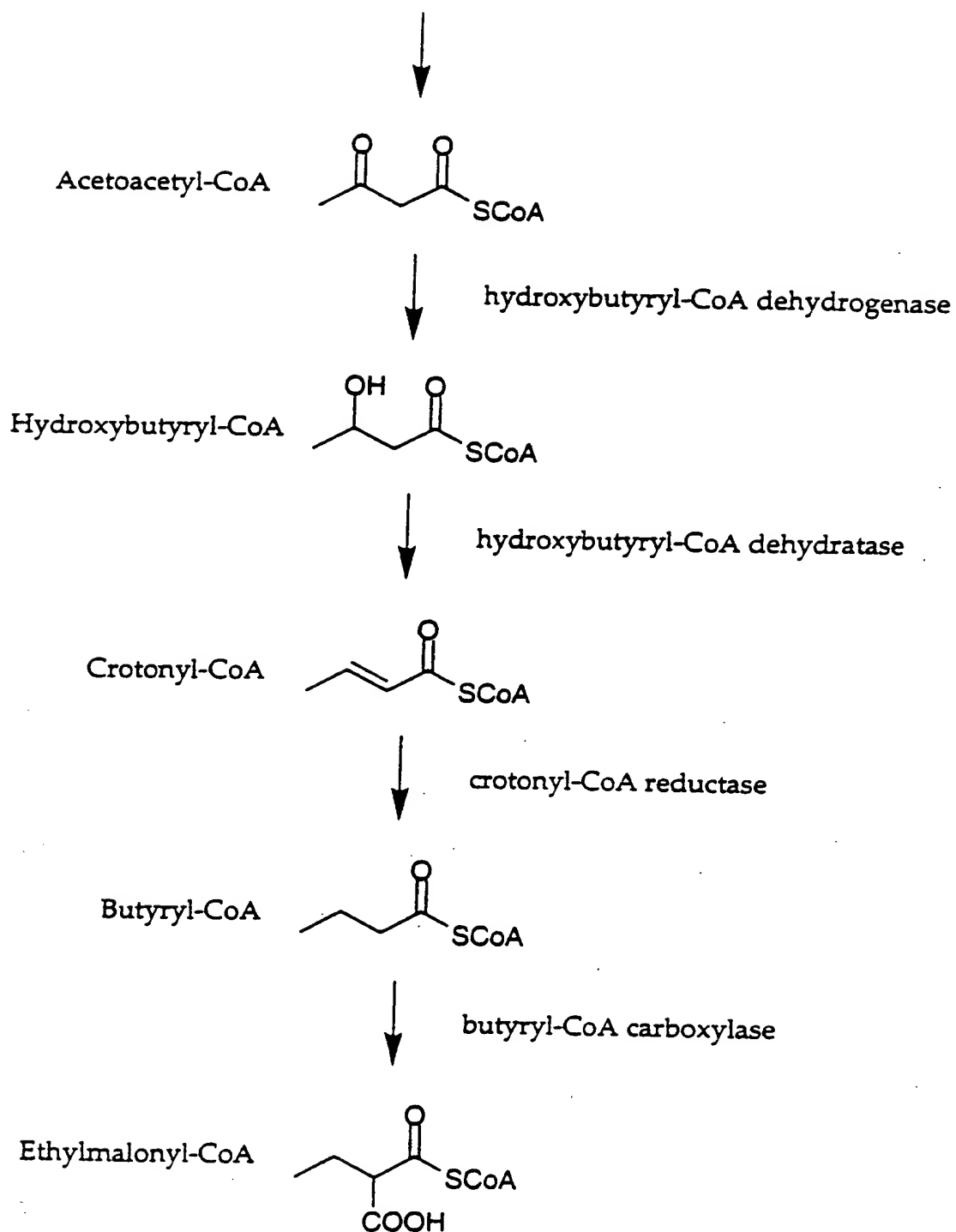


Figure 4

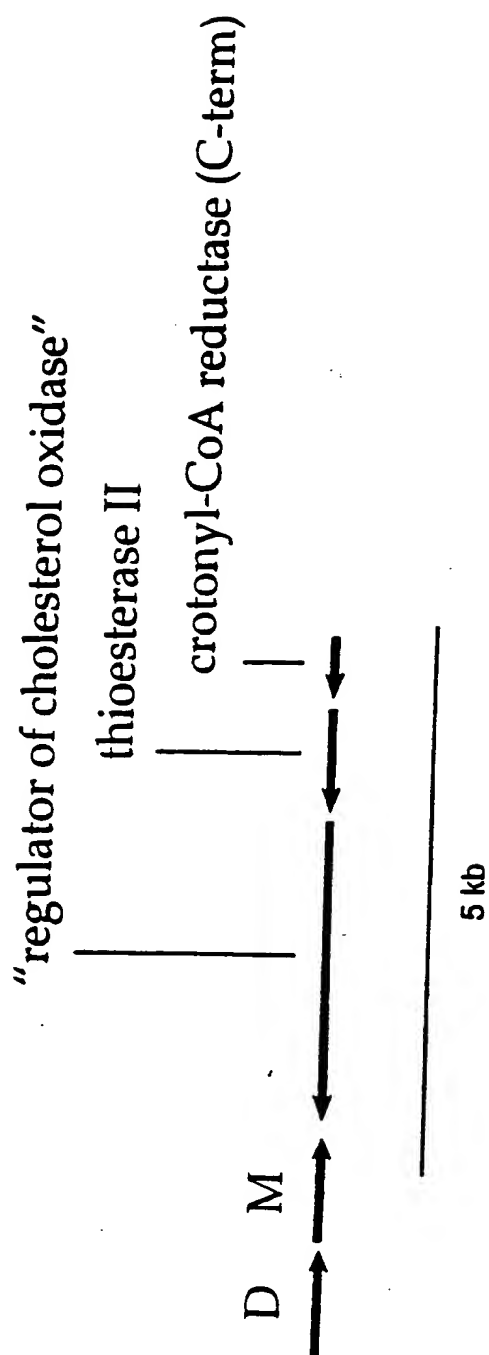


Figure 5

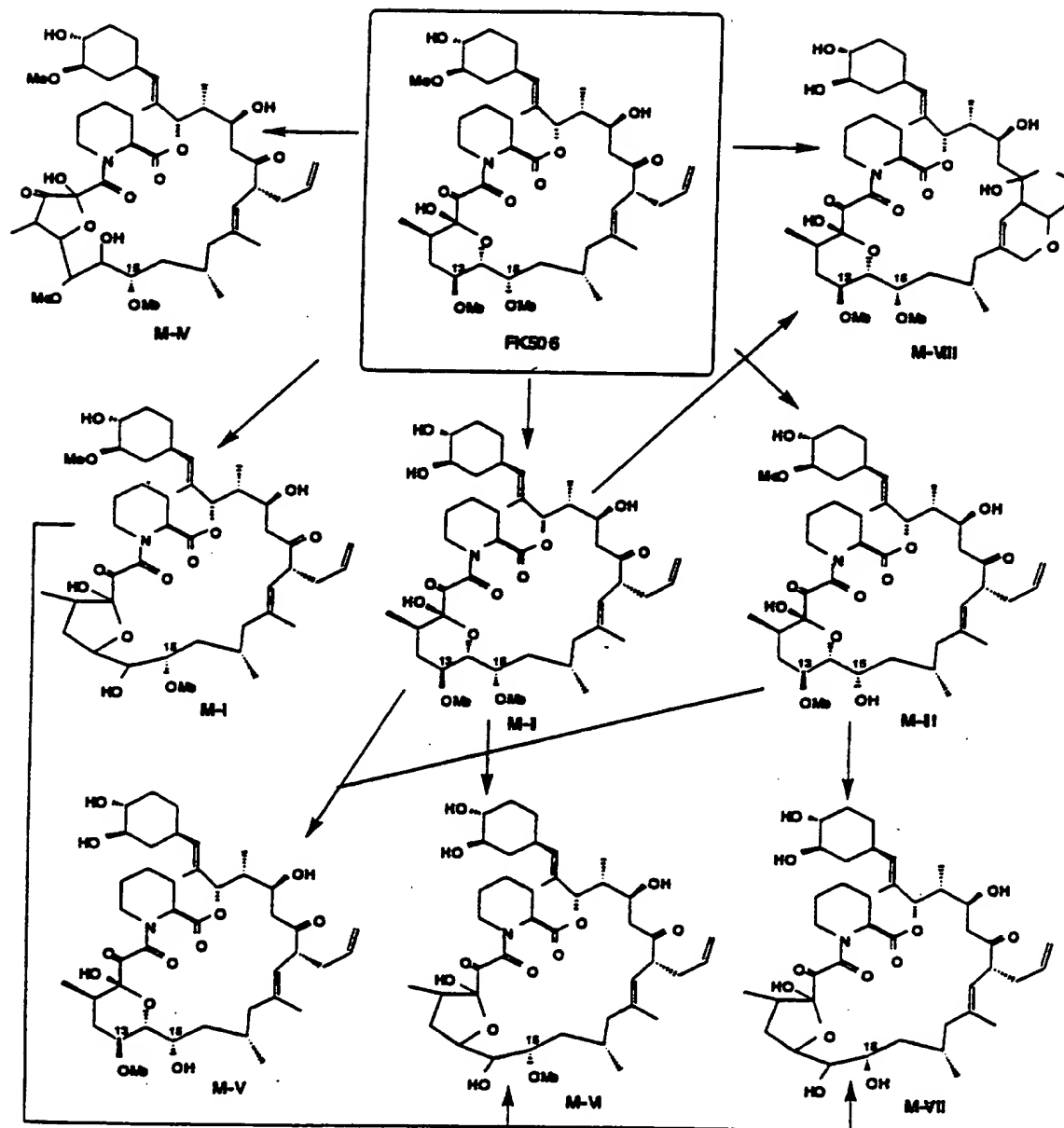


Figure 6

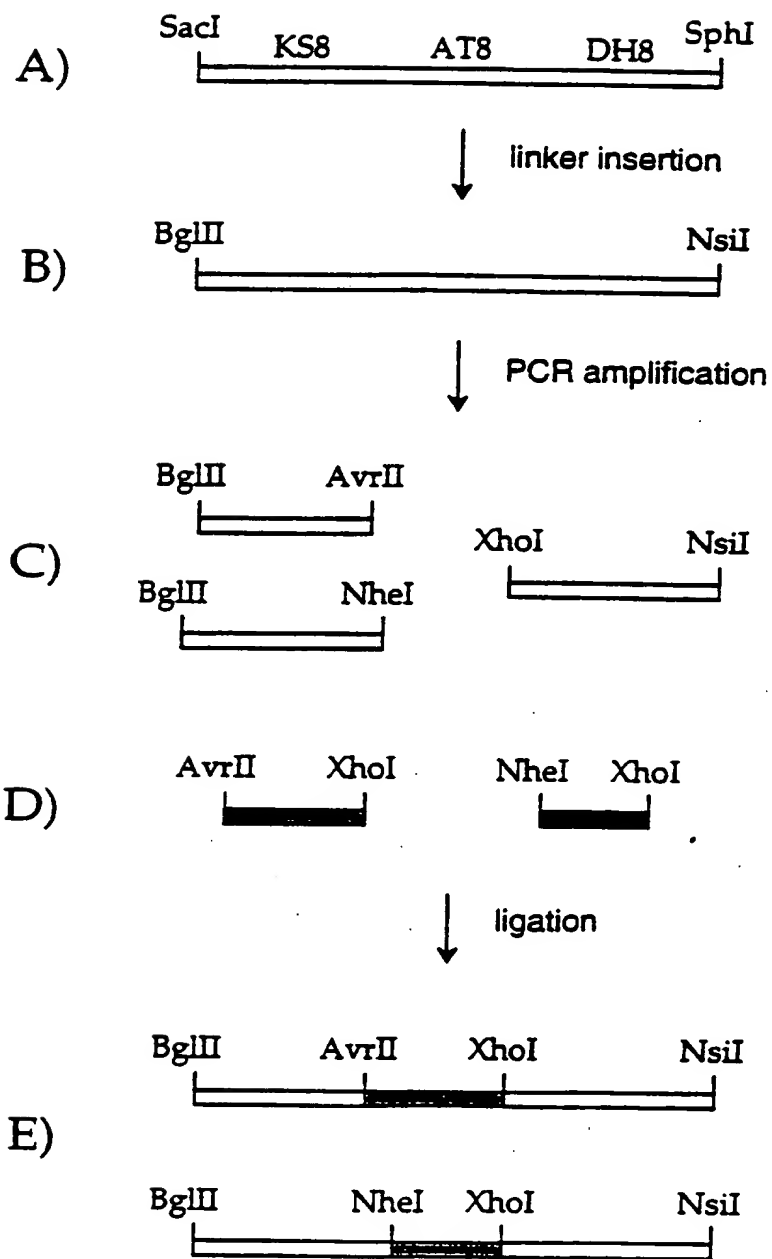


Figure 7

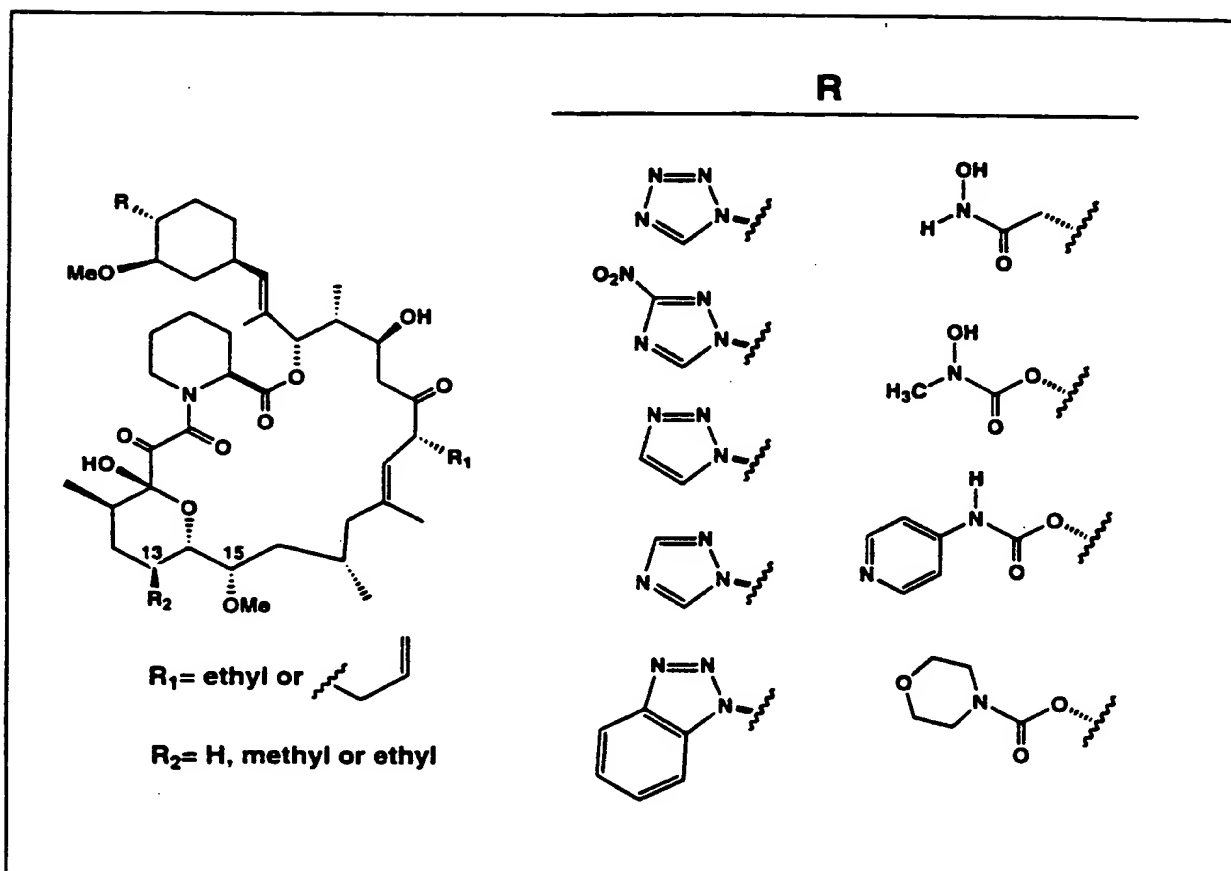


Figure 8
Part A





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ :C12N 15/52, 15/54, 15/62, 9/10, C12P
17/18, 19/32, C07D 498/18 // (C07D
498/18, 311:00, 273:00, 211:00)

A2

(11) International Publication Number:

WO 00/20601

(43) International Publication Date:

13 April 2000 (13.04.00)

(21) International Application Number: PCT/US99/22886

(22) International Filing Date: 1 October 1999 (01.10.99)

(30) Priority Data:

60/102,748	2 October 1998 (02.10.98)	US
60/123,810	11 March 1999 (11.03.99)	US
60/139,650	17 June 1999 (17.06.99)	US

(71) Applicant (for all designated States except US): KOSAN
BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive,
Hayward, CA 94545 (US).

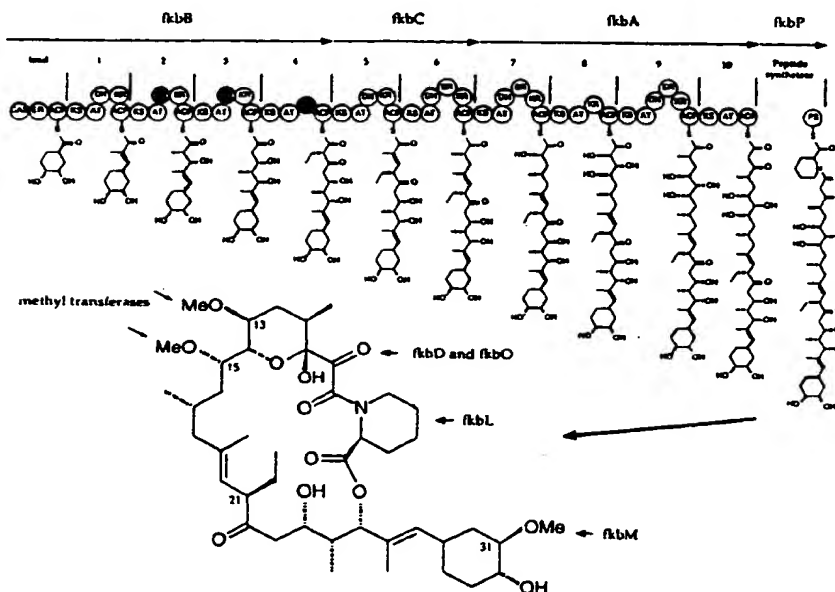
(72) Inventors; and

(75) Inventors/Applicants (for US only): REEVES, Christopher
[US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US).
CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA
95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue,
Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211
Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai
[CN/US]; 900 Constitution Drive, Foster City, CA 94404
(US).(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster
LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC
20006-1888 (US).(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN,
CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG,
KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX,
NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN,
ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ,
TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD,
RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK,
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI
patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR,
NE, SN, TD, TG).

Published

*Without international search report and to be republished
upon receipt of that report.**With an indication in relation to deposited biological
material furnished under Rule 13bis separately from the
description.*

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



(57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS
THEREFOR

5

Field of the Invention

The present invention relates to polyketides and the polyketide synthase (PKS) enzymes that produce them. The invention also relates generally to genes encoding PKS enzymes and to recombinant host cells containing such genes and in which expression of such genes leads to the production of polyketides. The present invention also relates to compounds useful as medicaments having immunosuppressive and/or neurotrophic activity. Thus, the invention relates to the fields of chemistry, molecular biology, and agricultural, medical, and veterinary technology.

10

Background of the Invention

15

Polyketides are a class of compounds synthesized from 2-carbon units through a series of condensations and subsequent modifications. Polyketides occur in many types of organisms, including fungi and mycelial bacteria, in particular, the actinomycetes.

Polyketides are biologically active molecules with a wide variety of structures, and the class encompasses numerous compounds with diverse activities. Tetracycline, erythromycin, epothilone, FK-506, FK-520, narbomycin, picromycin, rapamycin, spinocyn, and tylosin are examples of polyketides. Given the difficulty in producing polyketide compounds by traditional chemical methodology, and the typically low production of polyketides in wild-type cells, there has been considerable interest in finding improved or alternate means to produce polyketide compounds.

20

25

This interest has resulted in the cloning, analysis, and manipulation by recombinant DNA technology of genes that encode PKS enzymes. The resulting technology allows one to manipulate a known PKS gene cluster either to produce the polyketide synthesized by that PKS at higher levels than occur in nature or in hosts that otherwise do not produce the polyketide. The technology also allows one to produce molecules that are structurally related to, but distinct from, the polyketides produced from known PKS gene clusters. See, e.g., PCT publication Nos. WO 93/13663; 95/08548; 96/40968; 97/02358; 98/27203; and 98/49315; United States Patent Nos. 4,874,748; 5,063,155; 5,098,837; 5,149,639; 5,672,491; 5,712,146; 5,830,750; and 5,843,718; and Fu *et al.*, 1994, *Biochemistry* 33:

30

9321-9326; McDaniel *et al.*, 1993, *Science* 262: 1546-1550; and Rohr, 1995, *Angew. Chem. Int. Ed. Engl.* 34(8): 881-888, each of which is incorporated herein by reference.

Polyketides are synthesized in nature by PKS enzymes. These enzymes, which are complexes of multiple large proteins, are similar to the synthases that catalyze condensation of 2-carbon units in the biosynthesis of fatty acids. PKSs catalyze the biosynthesis of polyketides through repeated, decarboxylative Claisen condensations between acylthioester building blocks. The building blocks used to form complex polyketides are typically acylthioesters, such as acetyl, butyryl, propionyl, malonyl, hydroxymalonyl, methylmalonyl, and ethylmalonyl CoA. Other building blocks include amino acid like acylthioesters. PKS enzymes that incorporate such building blocks include an activity that functions as an amino acid ligase (an AMP ligase) or as a non-ribosomal peptide synthetase (NRPS). Two major types of PKS enzymes are known; these differ in their composition and mode of synthesis of the polyketide synthesized. These two major types of PKS enzymes are commonly referred to as Type I or "modular" and Type II "iterative" PKS enzymes.

In the Type I or modular PKS enzyme group, a set of separate catalytic active sites (each active site is termed a "domain", and a set thereof is termed a "module") exists for each cycle of carbon chain elongation and modification in the polyketide synthesis pathway. The typical modular PKS is composed of several large polypeptides, which can be segregated from amino to carboxy termini into a loading module, multiple extender modules, and a releasing (or thioesterase) domain. The PKS enzyme known as 6-deoxyerythronolide B synthase (DEBS) is a Type I PKS. In DEBS, there is a loading module, six extender modules, and a thioesterase (TE) domain. The loading module, six extender modules, and TE of DEBS are present on three separate proteins (designated DEBS-1, DEBS-2, and DEBS-3, with two extender modules per protein). Each of the DEBS polypeptides is encoded by a separate open reading frame (ORF) or gene; these genes are known as *eryAI*, *eryAII*, and *eryAIII*. See Caffrey *et al.*, 1992, *FEBS Letters* 304: 205, and U.S. Patent No. 5,824,513, each of which is incorporated herein by reference.

Generally, the loading module is responsible for binding the first building block used to synthesize the polyketide and transferring it to the first extender module. The loading module of DEBS consists of an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain. Another type of loading module utilizes an inactivated ketosynthase (KS) domain and AT and ACP domains. This inactivated KS is in some instances called KS^Q, where the superscript letter is the abbreviation for the amino acid, glutamine, that is

present instead of the active site cysteine required for ketosynthase activity. In other PKS enzymes, including the FK-506 PKS, the loading module incorporates an unusual starter unit and is composed of a CoA ligase like activity domain. In any event, the loading module recognizes a particular acyl-CoA (usually acetyl or propionyl but sometimes butyryl or other acyl-CoA) and transfers it as a thiol ester to the ACP of the loading module.

The AT on each of the extender modules recognizes a particular extender-CoA (malonyl or alpha-substituted malonyl, i.e., methylmalonyl, ethylmalonyl, and 2-hydroxymalonyl) and transfers it to the ACP of that extender module to form a thioester. Each extender module is responsible for accepting a compound from a prior module, binding a building block, attaching the building block to the compound from the prior module, optionally performing one or more additional functions, and transferring the resulting compound to the next module.

Each extender module of a modular PKS contains a KS, AT, ACP, and zero, one, two, or three domains that modify the beta-carbon of the growing polyketide chain. A typical (non-loading) minimal Type I PKS extender module is exemplified by extender module three of DEBS, which contains a KS domain, an AT domain, and an ACP domain. These three domains are sufficient to activate a 2-carbon extender unit and attach it to the growing polyketide molecule. The next extender module, in turn, is responsible for attaching the next building block and transferring the growing compound to the next extender module until synthesis is complete.

Once the PKS is primed with acyl- and malonyl-ACPs, the acyl group of the loading module is transferred to form a thiol ester (trans-esterification) at the KS of the first extender module; at this stage, extender module one possesses an acyl-KS and a malonyl (or substituted malonyl) ACP. The acyl group derived from the loading module is then covalently attached to the alpha-carbon of the malonyl group to form a carbon-carbon bond, driven by concomitant decarboxylation, and generating a new acyl-ACP that has a backbone two carbons longer than the loading building block (elongation or extension).

The polyketide chain, growing by two carbons each extender module, is sequentially passed as covalently bound thiol esters from extender module to extender module, in an assembly line-like process. The carbon chain produced by this process alone would possess a ketone at every other carbon atom, producing a polyketone, from which the name polyketide arises. Most commonly, however, additional enzymatic activities modify the beta

keto group of each two carbon unit just after it has been added to the growing polyketide chain but before it is transferred to the next module.

Thus, in addition to the minimal module containing KS, AT, and ACP domains necessary to form the carbon-carbon bond, and as noted above, other domains that modify the beta-carbonyl moiety can be present. Thus, modules may contain a ketoreductase (KR) domain that reduces the keto group to an alcohol. Modules may also contain a KR domain plus a dehydratase (DH) domain that dehydrates the alcohol to a double bond. Modules may also contain a KR domain, a DH domain, and an enoylreductase (ER) domain that converts the double bond product to a saturated single bond using the beta carbon as a methylene function. An extender module can also contain other enzymatic activities, such as, for example, a methylase or dimethylase activity.

After traversing the final extender module, the polyketide encounters a releasing domain that cleaves the polyketide from the PKS and typically cyclizes the polyketide. For example, final synthesis of 6-dEB is regulated by a TE domain located at the end of extender module six. In the synthesis of 6-dEB, the TE domain catalyzes cyclization of the macrolide ring by formation of an ester linkage. In FK-506, FK-520, rapamycin, and similar polyketides, the TE activity is replaced by a RapP (for rapamycin) or RapP like activity that makes a linkage incorporating a pipecolate acid residue. The enzymatic activity that catalyzes this incorporation for the rapamycin enzyme is known as RapP, encoded by the *rapP* gene. The polyketide can be modified further by tailoring enzymes; these enzymes add carbohydrate groups or methyl groups, or make other modifications, i.e., oxidation or reduction, on the polyketide core molecule. For example, 6-dEB is hydroxylated at C-6 and C-12 and glycosylated at C-3 and C-5 in the synthesis of erythromycin A.

In Type I PKS polypeptides, the order of catalytic domains is conserved. When all beta-keto processing domains are present in a module, the order of domains in that module from N-to-C-terminus is always KS, AT, DH, ER, KR, and ACP. Some or all of the beta-keto processing domains may be missing in particular modules, but the order of the domains present in a module remains the same. The order of domains within modules is believed to be important for proper folding of the PKS polypeptides into an active complex. Importantly, there is considerable flexibility in PKS enzymes, which allows for the genetic engineering of novel catalytic complexes. The engineering of these enzymes is achieved by modifying, adding, or deleting domains, or replacing them with those taken from other Type I PKS enzymes. It is also achieved by deleting, replacing, or adding entire modules with those

taken from other sources. A genetically engineered PKS complex should of course have the ability to catalyze the synthesis of the product predicted from the genetic alterations made.

Alignments of the many available amino acid sequences for Type I PKS enzymes has approximately defined the boundaries of the various catalytic domains. Sequence
5 alignments also have revealed linker regions between the catalytic domains and at the N- and C-termini of individual polypeptides. The sequences of these linker regions are less well conserved than are those for the catalytic domains, which is in part how linker regions are identified. Linker regions can be important for proper association between domains and between the individual polypeptides that comprise the PKS complex. One can thus view the
10 linkers and domains together as creating a scaffold on which the domains and modules are positioned in the correct orientation to be active. This organization and positioning, if retained, permits PKS domains of different or identical substrate specificities to be substituted (usually at the DNA level) between PKS enzymes by various available methodologies. In selecting the boundaries of, for example, an AT replacement, one can
15 thus make the replacement so as to retain the linkers of the recipient PKS or to replace them with the linkers of the donor PKS AT domain, or, preferably, make both constructs to ensure that the correct linker regions between the KS and AT domains have been included in at least one of the engineered enzymes. Thus, there is considerable flexibility in the design of new PKS enzymes with the result that known polyketides can be produced more
20 effectively, and novel polyketides useful as pharmaceuticals or for other purposes can be made.

By appropriate application of recombinant DNA technology, a wide variety of polyketides can be prepared in a variety of different host cells provided one has access to nucleic acid compounds that encode PKS proteins and polyketide modification enzymes.
25 The present invention helps meet the need for such nucleic acid compounds by providing recombinant vectors that encode the FK-520 PKS enzyme and various FK-520 modification enzymes. Moreover, while the FK-506 and FK-520 polyketides have many useful activities, there remains a need for compounds with similar useful activities but with better pharmacokinetic profile and metabolism and fewer side-effects. The present invention helps
30 meet the need for such compounds as well.

Summary of the Invention

In one embodiment, the present invention provides recombinant DNA vectors that encode all or part of the FK-520 PKS enzyme. Illustrative vectors of the invention include cosmid pKOS034-120, pKOS034-124, pKOS065-C31, pKOS065-C3, pKOS065-M27, and pKOS065-M21. The invention also provides nucleic acid compounds that encode the various domains of the FK-520 PKS, i.e., the KS, AT, ACP, KR, DH, and ER domains. These compounds can be readily used, alone or in combination with nucleic acids encoding other FK-520 or non-FK-520 PKS domains, as intermediates in the construction of recombinant vectors that encode all or part of PKS enzymes that make novel polyketides.

The invention also provides isolated nucleic acids that encode all or part of one or more modules of the FK-520 PKS, each module comprising a ketosynthase activity, an acyl transferase activity, and an acyl carrier protein activity. The invention provides an isolated nucleic acid that encodes one or more open reading frames of FK-520 PKS genes, said open reading frames comprising coding sequences for a CoA ligase activity, an NRPS activity, or two or more extender modules. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides isolated nucleic acids that encode all or a part of a PKS that contains at least one module in which at least one of the domains in the module is a domain from a non-FK-520 PKS and at least one domain is from the FK-520 PKS. The non-FK-520 PKS domain or module originates from the rapamycin PKS, the FK-506 PKS, DEBS, or another PKS. The invention also provides recombinant expression vectors containing these nucleic acids.

In another embodiment, the invention provides a method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector that encodes at least one module of a PKS, said module comprising at least one FK-520 PKS domain, and culturing said host cell under conditions such that said PKS is produced and catalyzes synthesis of said polyketide. In one aspect, the method is practiced with a *Streptomyces* host cell. In another aspect, the polyketide produced is FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-520. In another aspect, the polyketide produced is a polyketide related in structure to FK-506 or rapamycin.

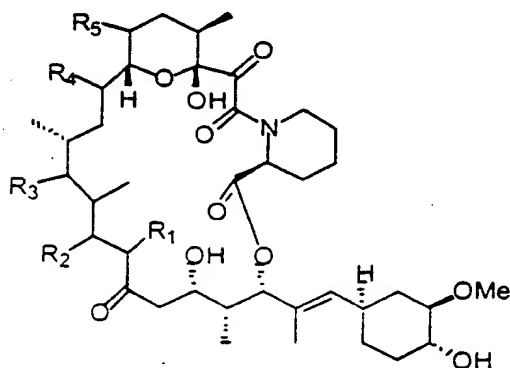
In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of ethylmalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require ethylmalonyl CoA for biosynthesis.

The invention also provides recombinant nucleic acids that encode AT domains specific for ethylmalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring ethylmalonyl CoA in host cells that otherwise are unable to produce such polyketides.

5 In another embodiment, the invention provides a set of genes in recombinant form sufficient for the synthesis of 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA in a heterologous host cell. These genes and the methods of the invention enable one to create recombinant host cells with the ability to produce polyketides or other compounds that require 2-hydroxymalonyl CoA for biosynthesis. The invention also provides recombinant
10 nucleic acids that encode AT domains specific for 2-hydroxymalonyl CoA and 2-methoxymalonyl CoA. Thus, the compounds of the invention can be used to produce polyketides requiring 2-hydroxymalonyl CoA or 2-methoxymalonyl CoA in host cells that are otherwise unable to produce such polyketides.

In another embodiment, the invention provides a compound related in structure to
15 FK-520 or FK-506 that is useful in the treatment of a medical condition. These compounds include compounds in which the C-13 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. Such compounds are less susceptible to the main *in vivo* pathway of degradation for FK-520 and FK-506 and related compounds and thus exhibit an improved pharmacokinetic profile. The compounds of the
20 invention also include compounds in which the C-15 methoxy group is replaced by a moiety selected from the group consisting of hydrogen, methyl, and ethyl moieties. The compounds of the invention also include the above compounds further modified by chemical methodology to produce derivatives such as, but not limited to, the C-18 hydroxyl derivatives, which have potent neurotrophin but not immunosuppression activities.

25 Thus, the invention provides polyketides having the structure:



wherein, R_1 is hydrogen, methyl, ethyl, or allyl; R_2 is hydrogen or hydroxyl, provided that when R_2 is hydrogen, there is a double bond between C-20 and C-19; R_3 is hydrogen or hydroxyl; R_4 is methoxyl, hydrogen, methyl, or ethyl; and R_5 is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506. The invention provides these compounds in purified form and in pharmaceutical compositions.

In another embodiment, the invention provides a method for treating a medical condition by administering a pharmaceutically efficacious dose of a compound of the invention. The compounds of the invention may be administered to achieve immunosuppression or to stimulate nerve growth and regeneration.

These and other embodiments and aspects of the invention will be more fully understood after consideration of the attached Drawings and their brief description below, together with the detailed description, examples, and claims that follow.

Brief Description of the Drawings

Figure 1 shows a diagram of the FK-520 biosynthetic gene cluster. The top line provides a scale in kilobase pairs (kb). The second line shows a restriction map with selected restriction enzyme recognition sequences indicated. K is *KpnI*; X is *XhoI*, S is *SacI*; P is *PstI*; and E is *EcoRI*. The third line indicates the position of FK-520 PKS and related genes. Genes are abbreviated with a one letter designation, i.e., C is *fkbc*. Immediately under the third line are numbered segments showing where the loading module (L) and ten different extender modules (numbered 1 - 10) are encoded on the various genes shown. At the bottom of the Figure, the DNA inserts of various cosmids of the invention (i.e., 34-124 is cosmid pKOS034-124) are shown in alignment with the FK-520 biosynthetic gene cluster.

Figure 2 shows the loading module (load), the ten extender modules, and the peptide synthetase domain of the FK-520 PKS, together with, on the top line, the genes that encode the various domains and modules. Also shown are the various intermediates in FK-520 biosynthesis, as well as the structure of FK-520, with carbons 13, 15, 21, and 31 numbered. The various domains of each module and subdomains of the loading module are also shown. The darkened circles showing the DH domains in modules 2, 3, and 4 indicate that the dehydratase domain is not functional as a dehydratase; this domain may affect the

stereochemistry at the corresponding position in the polyketide. The substituents on the FK-520 structure that result from the action of non-PKS enzymes are also indicated by arrows, together with the types of enzymes or the genes that code for the enzymes that mediate the action. Although the methyltransferase is shown acting at the C-13 and C-15 hydroxyl groups after release of the polyketide from the PKS, the methyltransferase may act on the 2-hydroxymalonyl substrate prior to or contemporaneously with its incorporation during polyketide synthesis.

Figure 3 shows a close-up view of the left end of the FK-520 gene cluster, which contains at least ten additional genes. The ethyl side chain on carbon 21 of FK-520 (Figure 2) is derived from an ethylmalonyl CoA extender unit that is incorporated by an ethylmalonyl specific AT domain in extender module 4 of the PKS. At least four of the genes in this region code for enzymes involved in ethylmalonyl biosynthesis. The polyhydroxybutyrate depolymerase is involved in maintaining hydroxybutyryl-CoA pools during FK-520 production. Polyhydroxybutyrate accumulates during vegetative growth and disappears during stationary phase in other *Streptomyces* (Ranade and Vining, 1993, *Can. J. Microbiol.* 39:377). Open reading frames with unknown function are indicated with a question mark.

Figure 4 shows a biosynthetic pathway for the biosynthesis of ethylmalonyl CoA from acetoacetyl CoA consistent with the function assigned to four of the genes in the FK-520 gene cluster shown in Figure 3.

Figure 5 shows a close-up view of the right-end of the FK-520 PKS gene cluster (and of the sequences on cosmid pKOS065-C31). The genes shown include *fk bD*, *fk bM* (a methyl transferase that methylates the hydroxyl group on C-31 of FK-520), *fk bN* (a homolog of a gene described as a regulator of cholesterol oxidase and that is believed to be a transcriptional activator), *fk bQ* (a type II thioesterase, which can increase polyketide production levels), and *fk bS* (a crotonyl-CoA reductase involved in the biosynthesis of ethylmalonyl CoA).

Figure 6 shows the proposed degradative pathway for tacrolimus (FK-506) metabolism.

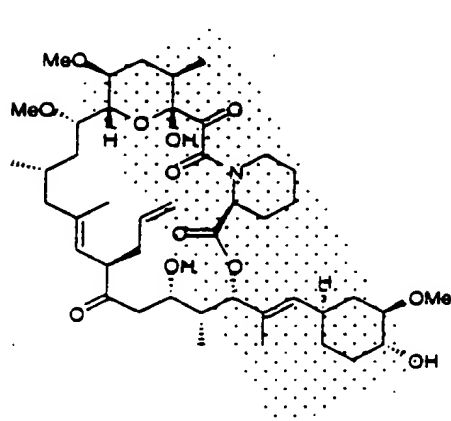
Figure 7 shows a schematic process for the construction of recombinant PKS genes of the invention that encode PKS enzymes that produce 13-desmethoxy FK-506 and FK-520 polyketides of the invention, as described in Example 4, below.

Figure 8, in Parts A and B, shows certain compounds of the invention preferred for dermal application in Part A and a synthetic route for making those compounds in Part B.

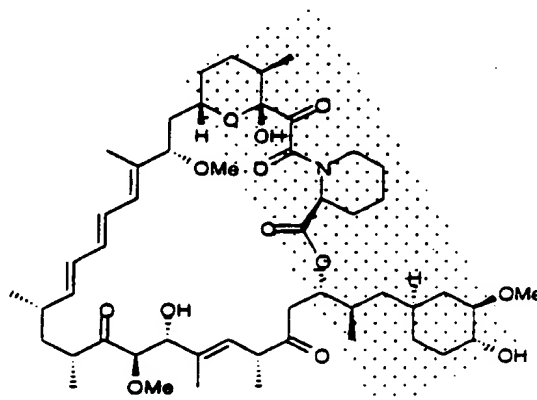
Detailed Description of the Invention

5 Given the valuable pharmaceutical properties of polyketides, there is a need for methods and reagents for producing large quantities of polyketides, as well as for producing related compounds not found in nature. The present invention provides such methods and reagents, with particular application to methods and reagents for producing the polyketides known as FK-520, also known as ascomycin or L-683,590 (see Holt *et al.*, 1993, *JACS* 10 115:9925), and FK-506, also known as tacrolimus. Tacrolimus is a macrolide immunosuppressant used to prevent or treat rejection of transplanted heart, kidney, liver, lung, pancreas, and small bowel allografts. The drug is also useful for the prevention and treatment of graft-versus-host disease in patients receiving bone marrow transplants, and for the treatment of severe, refractory uveitis. There have been additional reports of the 15 unapproved use of tacrolimus for other conditions, including alopecia universalis, autoimmune chronic active hepatitis, inflammatory bowel disease, multiple sclerosis, primary biliary cirrhosis, and scleroderma. The invention provides methods and reagents for making novel polyketides related in structure to FK-520 and FK-506, and structurally related polyketides such as rapamycin.

20 The FK-506 and rapamycin polyketides are potent immunosuppressants, with chemical structures shown below.



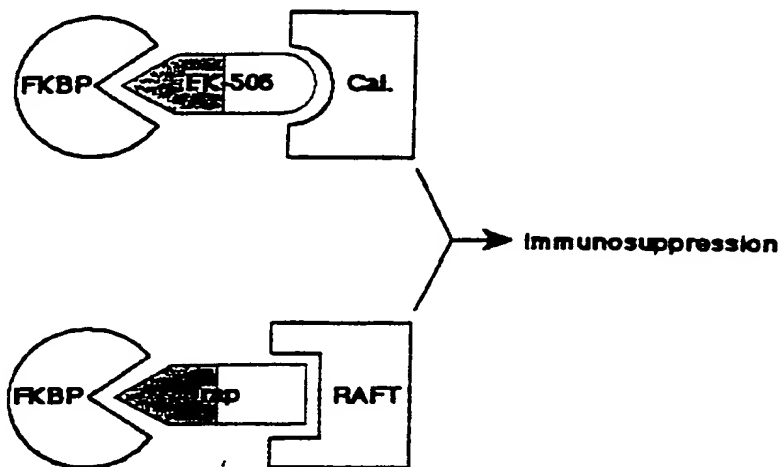
FK-506



Rapamycin

FK-520 differs from FK-506 in that it lacks the allyl group at C-21 of FK-506, having instead an ethyl group at that position, and has similar activity to FK-506, albeit reduced immunosuppressive activity.

These compounds act through initial formation of an intermediate complex with protein "immunophilins" known as FKBP (FK-506 binding proteins), including FKBP-12. Immunophilins are a class of cytosolic proteins that form complexes with molecules such as FK-506, FK-520, and rapamycin that in turn serve as ligands for other cellular targets involved in signal transduction. Binding of FK-506, FK-520, and rapamycin to FKBP occurs through the structurally similar segments of the polyketide molecules, known as the "FKBP-binding domain" (as generally but not precisely indicated by the stippled regions in the structures above). The FK-506-FKBP complex then binds calcineurin, while the rapamycin-FKBP complex binds to a protein known as RAFT-1. Binding of the FKBP-polyketide complex to these second proteins occurs through the dissimilar regions of the drugs known as the "effector" domains.



15

The three component FKBP-polyketide-effector complex is required for signal transduction and subsequent immunosuppressive activity of FK-506, FK-520, and rapamycin. Modifications in the effector domains of FK-506, FK-520, and rapamycin that destroy binding to the effector proteins (calcineurin or RAFT) lead to loss of immunosuppressive activity, even though FKBP binding is unaffected. Further, such analogs antagonize the immunosuppressive effects of the parent polyketides, because they compete for FKBP. Such non-immunosuppressive analogs also show reduced toxicity (see Dumont *et al.*, 1992, *Journal of Experimental Medicine* 176, 751-760), indicating that much of the toxicity of these drugs is not linked to FKBP binding.

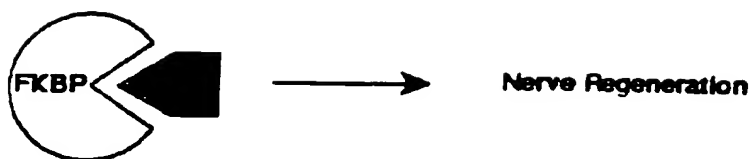
20

In addition to immunosuppressive activity, FK-520, FK-506, and rapamycin have neurotrophic activity. In the central nervous system and in peripheral nerves, immunophilins are referred to as "neuroimmunophilins". The neuroimmunophilin FKBP is markedly enriched in the central nervous system and in peripheral nerves. Molecules that bind to the neuroimmunophilin FKBP, such as FK-506 and FK-520, have the remarkable effect of stimulating nerve growth. *In vitro*, they act as neurotrophins, i.e., they promote neurite outgrowth in NGF-treated PC12 cells and in sensory neuronal cultures, and in intact animals, they promote regrowth of damaged facial and sciatic nerves, and repair lesioned serotonin and dopamine neurons in the brain. See Gold *et al.*, Jun. 1999, *J. Pharm. Exp. Ther.* 289(3): 1202-1210; Lyons *et al.*, 1994, *Proc. National Academy of Science* 91: 3191-3195; Gold *et al.*, 1995, *Journal of Neuroscience* 15: 7509-7516; and Steiner *et al.*, 1997, *Proc. National Academy of Science* 94: 2019-2024. Further, the restored central and peripheral neurons appear to be functional.

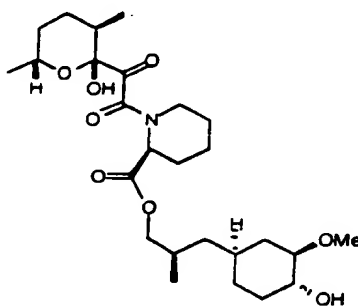
Compared to protein neurotrophic molecules (BDNF, NGF, etc.), the small-molecule neurotrophins such as FK-506, FK-520, and rapamycin have different, and often advantageous, properties. First, whereas protein neurotrophins are difficult to deliver to their intended site of action and may require intra-cranial injection, the small-molecule neurotrophins display excellent bioavailability; they are active when administered subcutaneously and orally. Second, whereas protein neurotrophins show quite specific effects, the small-molecule neurotrophins show rather broad effects. Finally, whereas protein neurotrophins often show effects on normal sensory nerves, the small-molecule neurotrophins do not induce aberrant sprouting of normal neuronal processes and seem to affect damaged nerves specifically. Neuroimmunophilin ligands have potential therapeutic utility in a variety of disorders involving nerve degeneration (e.g. multiple sclerosis, Parkinson's disease, Alzheimer's disease, stroke, traumatic spinal cord and brain injury, peripheral neuropathies).

Recent studies have shown that the immunosuppressive and neurite outgrowth activity of FK-506, FK-520, and rapamycin can be separated; the neuroregenerative activity in the absence of immunosuppressive activity is retained by agents which bind to FKBP but not to the effector proteins calcineurin or RAFT. See Steiner *et al.*, 1997, *Nature Medicine* 3: 421-428.

13



Available structure-activity data show that the important features for neurotrophic activity of rapamycin, FK-520, and FK-506 lie within the common, contiguous segments of the macrolide ring that bind to FKBP. This portion of the molecule is termed the "FKBP binding domain" (see VanDuyne *et al.*, 1993, *Journal of Molecular Biology* 229: 105-124.). Nevertheless, the effector domains of the parent macrolides contribute to conformational rigidity of the binding domain and thus indirectly contribute to FKBP binding.

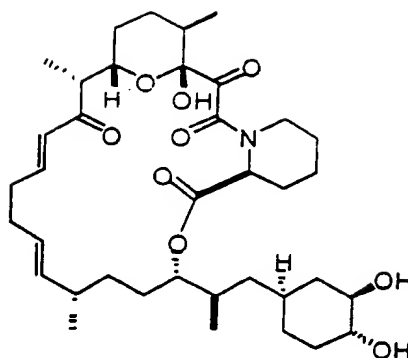


"FKBP binding domain"

There are a number of other reported analogs of FK-506, FK-520, and rapamycin that bind to FKBP but not the effector protein calcineurin or RAFT. These analogs show effects on nerve regeneration without immunosuppressive effects.

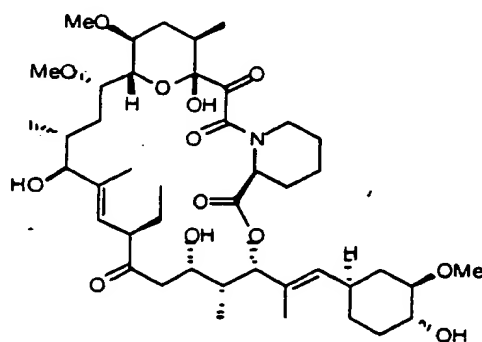
Naturally occurring FK-520 and FK-506 analogs include the antascomycins, which are FK-506-like macrolides that lack the functional groups of FK-506 that bind to calcineurin (see Fehr *et al.*, 1996, *The Journal of Antibiotics* 49: 230-233). These molecules bind FKBP as effectively as does FK-506; they antagonize the effects of both FK-506 and rapamycin, yet lack immunosuppressive activity.

14

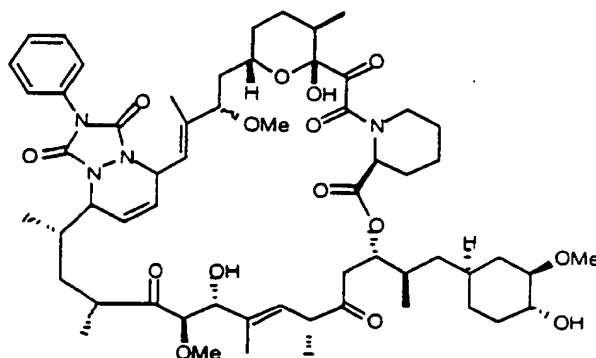


Antascomycin A

Other analogs can be produced by chemically modifying FK-506, FK-520, or rapamycin. One approach to obtaining neuroimmunophilin ligands is to destroy the effector binding region of FK-506, FK-520, or rapamycin by chemical modification. While the chemical modifications permitted on the parent compounds are quite limited, some useful chemically modified analogs exist. The FK-520 analog L-685,818 ($ED_{50} = 0.7$ nM for FKBP binding; see Dumont *et al.*, 1992), and the rapamycin analog WAY-124,466 ($IC_{50} = 12.5$ nM; see Ocain *et al.*, 1993, *Biochemistry Biophysical Research Communications* 192: 1340-134693) are about as effective as FK-506, FK-520, and rapamycin at promoting neurite outgrowth in sensory neurons (see Steiner *et al.*, 1997).



L-685,818

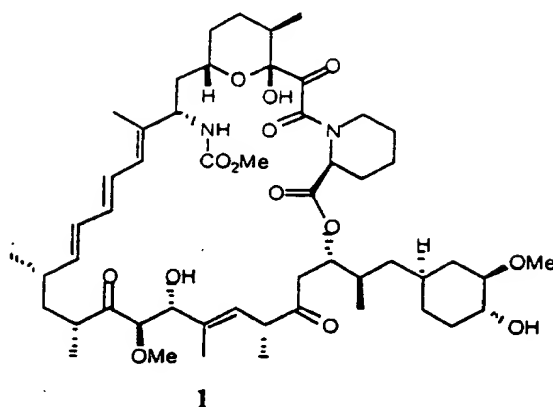


WAY-124,466

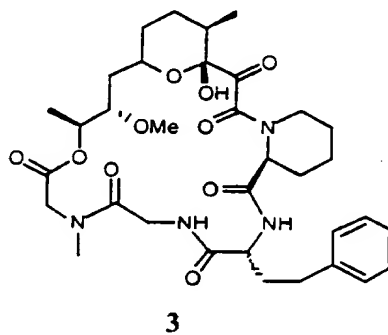
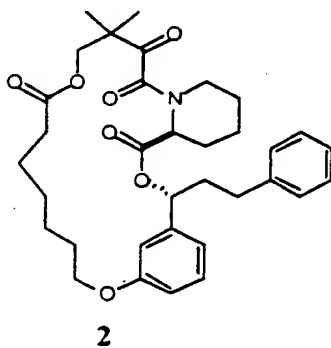
One of the few positions of rapamycin that is readily amenable to chemical modification is the allylic 16-methoxy group; this reactive group is readily exchanged by acid-catalyzed nucleophilic substitution. Replacement of the 16-methoxy group of rapamycin with a variety of bulky groups has produced analogs showing selective loss of immunosuppressive activity while retaining FKBP-binding (see Luengo *et al.*, 1995, *Chemistry & Biology* 2: 471-481). One of the best compounds, 1, below, shows complete

SUBSTITUTE SHEET (RULE 26)

loss of activity in the splenocyte proliferation assay with only a 10-fold reduction in binding to FKBP.



There are also synthetic analogs of FKBP binding domains. These compounds reflect an approach to obtaining neuroimmunophilin ligands based on "rationally designed" molecules that retain the FKBP-binding region in an appropriate conformation for binding to FKBP, but do not possess the effector binding regions. In one example, the ends of the FKBP binding domain were tethered by hydrocarbon chains (see Holt *et al.*, 1993, *Journal of the American Chemical Society* 115: 9925-9938); the best analog, 2, below, binds to FKBP about as well as FK-506. In a similar approach, the ends of the FKBP binding domain were tethered by a tripeptide to give analog 3, below, which binds to FKBP about 20-fold poorer than FK-506. These compounds are anticipated to have neuroimmunophilin binding activity.



15

In a primate MPTP model of Parkinson's disease, administration of FKBP ligand GPI-1046 caused brain cells to regenerate and behavioral measures to improve. MPTP is a neurotoxin, which, when administered to animals, selectively damages nigral-striatal dopamine neurons in the brain, mimicking the damage caused by Parkinson's disease. Whereas, before treatment, animals were unable to use affected limbs, the FKBP ligand

20

restored the ability of animals to feed themselves and gave improvements in measures of locomotor activity, neurological outcome, and fine motor control. There were also corresponding increases in regrowth of damaged nerve terminals. These results demonstrate the utility of FKBP ligands for treatment of diseases of the CNS.

5 From the above description, two general approaches towards the design of non-immunosuppressant, neuroimmunophilin ligands can be seen. The first involves the construction of constrained cyclic analogs of FK-506 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP. The advantages of this approach are that the conformation of the analogs can be accurately modeled and predicted by
10 computational methods, and the analogs closely resemble parent molecules that have proven pharmacological properties. A disadvantage is that the difficult chemistry limits the numbers and types of compounds that can be prepared. The second approach involves the trial and error construction of acyclic analogs of the FKBP binding domain by conventional medicinal chemistry. The advantages to this approach are that the chemistry is suitable for
15 production of the numerous compounds needed for such interactive chemistry-bioassay approaches. The disadvantages are that the molecular types of compounds that have emerged have no known history of appropriate pharmacological properties, have rather labile ester functional groups, and are too conformationally mobile to allow accurate prediction of conformational properties.

20 The present invention provides useful methods and reagents related to the first approach, but with significant advantages. The invention provides recombinant PKS genes that produce a wide variety of polyketides that cannot otherwise be readily synthesized by chemical methodology alone. Moreover, the present invention provides polyketides that have either or both of the desired immunosuppressive and neurotrophic activities, some of
25 which are produced only by fermentation and others of which are produced by fermentation and chemical modification. Thus, in one aspect, the invention provides compounds that optimally bind to FKBP but do not bind to the effector proteins. The methods and reagents of the invention can be used to prepare numerous constrained cyclic analogs of FK-520 in which the FKBP binding domain is fixed in a conformation optimal for binding to FKBP.
30 Such compounds will show neuroimmunophilin binding (neurotrophic) but not immunosuppressive effects. The invention also allows direct manipulation of FK-520 and related chemical structures *via* genetic engineering of the enzymes involved in the biosynthesis of FK-520 (as well as related compounds, such as FK-506 and rapamycin);

similar chemical modifications are simply not possible because of the complexity of the structures. The invention can also be used to introduce "chemical handles" into normally inert positions that permit subsequent chemical modifications.

5 Several general approaches to achieve the development of novel neuroimmunophilin ligands are facilitated by the methods and reagents of the present invention. One approach is to make "point mutations" of the functional groups of the parent FK-520 structure that bind to the effector molecules to eliminate their binding potential. These types of structural modifications are difficult to perform by chemical modification, but can be readily accomplished with the methods and reagents of the invention.

10 A second, more extensive approach facilitated by the present invention is to utilize molecular modeling to predict optimal structures *ab initio* that bind to FKBP but not effector molecules. Using the available X-ray crystal structure of FK-520 (or FK-506) bound to FKBP, molecular modeling can be used to predict polyketides that should optimally bind to FKBP but not calcineurin. Various macrolide structures can be generated
15 by linking the ends of the FKBP-binding domain with "all possible" polyketide chains of variable length and substitution patterns that can be prepared by genetic manipulation of the FK-520 or FK-506 PKS gene cluster in accordance with the methods of the invention. The ground state conformations of the virtual library can be determined, and compounds that possess binding domains most likely to bind well to FKBP can be prepared and tested.

20 Once a compound is identified in accordance with the above approaches, the invention can be used to generate a focused library of analogs around the lead candidate, to "fine tune" the compound for optimal properties. Finally, the genetic engineering methods of the invention can be directed towards producing "chemical handles" that enable medicinal chemists to modify positions of the molecule previously inert to chemical
25 modification. This opens the path to previously prohibited chemical optimization of lead compounds by time-proven approaches.

Moreover, the present invention provides polyketide compounds and the recombinant genes for the PKS enzymes that produce the compounds that have significant advantages over FK-506 and FK-520 and their analogs. The metabolism and
30 pharmacokinetics of tacrolimus has been extensively studied, and FK-520 is believed to be similar in these respects. Absorption of tacrolimus is rapid, variable, and incomplete from the gastrointestinal tract (Harrison's Principles of Internal Medicine, 14th edition, 1998, McGraw Hill, 14, 20, 21, 64-67). The mean bioavailability of the oral dosage form is 27%.

(range 5 to 65%). The volume of distribution (VolD) based on plasma is 5 to 65 L per kg of body weight (L/kg), and is much higher than the VolD based on whole blood concentrations, the difference reflecting the binding of tacrolimus to red blood cells. Whole blood concentrations may be 12 to 67 times the plasma concentrations. Protein binding is high (75 to 99%), primarily to albumin and alpha₁-acid glycoprotein. The half-life for distribution is 0.9 hour; elimination is biphasic and variable: terminal-11.3 hr (range, 3.5 to 40.5 hours). The time to peak concentration is 0.5 to 4 hours after oral administration.

Tacrolimus is metabolized primarily by cytochrome P450 3A enzymes in the liver and small intestine. The drug is extensively metabolized with less than 1% excreted unchanged in urine. Because hepatic dysfunction decreases clearance of tacrolimus, doses have to be reduced substantially in primary graft non-function, especially in children. In addition, drugs that induce the cytochrome P450 3A enzymes reduce tacrolimus levels, while drugs that inhibit these P450s increase tacrolimus levels. Tacrolimus bioavailability doubles with co-administration of ketoconazole, a drug that inhibits P450 3A. See, Vincent *et al.*, 1992, *In vitro* metabolism of FK-506 in rat, rabbit, and human liver microsomes: Identification of a major metabolite and of cytochrome P450 3A as the major enzymes responsible for its metabolism, *Arch. Biochem. Biophys.* 294: 454-460; Iwasaki *et al.*, 1993, Isolation, identification, and biological activities of oxidative metabolites of FK-506, a potent immunosuppressive macrolide lactone, *Drug Metabolism & Disposition* 21: 971-977; Shiraga *et al.*, 1994, Metabolism of FK-506, a potent immunosuppressive agent, by cytochrome P450 3A enzymes in rat, dog, and human liver microsomes, *Biochem. Pharmacol.* 47: 727-735; and Iwasaki *et al.*, 1995, Further metabolism of FK-506 (Tacrolimus); Identification and biological activities of the metabolites oxidized at multiple sites of FK-506, *Drug Metabolism & Disposition* 23: 28-34. The cytochrome P450 3A subfamily of isozymes has been implicated as important in this degradative process.

Structures of the eight isolated metabolites formed by liver microsomes are shown in Figure 6. Four metabolites of FK-506 involve demethylation of the oxygens on carbons 13, 15, and 31, and hydroxylation of carbon 12. The 13-demethylated (hydroxy) compounds undergo cyclizations of the 13-hydroxy at C-10 to give MI, MVI and MVII, and the 12-hydroxy metabolite at C-10 to give I. Another four metabolites formed by oxidation of the four metabolites mentioned above were isolated by liver microsomes from dexamethasone treated rats. Three of these are metabolites doubly demethylated at the methoxy groups on carbons 15 and 31 (M-V), 13 and 31 (M-VI), and 13 and 15 (M-VII). The fourth, M-VIII.

was the metabolite produced after demethylation of the 31-methoxy group, followed by formation of a fused ring system by further oxidation. Among the eight metabolites, M-II has immunosuppressive activity comparable to that of FK-506, whereas the other metabolites exhibit weak or negligible activities. Importantly, the major metabolite of human, dog, and rat liver microsomes is the 13-demethylated and cyclized FK-506 (M-I).

Thus, the major metabolism of FK-506 proceeds via 13-demethylation followed by cyclization to the inactive M-I, this representing about 90% of the metabolic products after a 10 minute incubation with liver microsomes. Analogs of tacrolimus that do not possess a C-13 methoxy group would not be susceptible to the first and most important biotransformation in the destructive metabolism of tacrolimus (i.e. cyclization of 13-hydroxy to C-10). Thus, a 13-desmethoxy analog of FK-506 should have a longer half-life in the body than does FK-506. The C-13 methoxy group is believed not to be required for binding to FKBP or calcineurin. The C-13 methoxy is not present on the identical position of rapamycin, which binds to FKBP with equipotent affinity as tacrolimus. Also, analysis of the 3-dimensional structure of the FKBP-tacrolimus-calcineurin complex shows that the C-13 methoxy has no interaction with FKBP and only a minor interaction with calcineurin. The present invention provides C-13-desmethoxy analogs of FK-506 and FK-520, as well as the recombinant genes that encode the PKS enzymes that catalyze their synthesis and host cells that produce the compounds.

These compounds exhibit, relative to their naturally occurring counterparts, prolonged immunosuppressive action *in vivo*, thereby allowing a lower dosage and/or reduced frequency of administration. Dosing is more predictable, because the variability in FK-506 dosage is largely due to variation of metabolism rate. FK-506 levels in blood can vary widely depending on interactions with drugs that induce or inhibit cytochrome P450 3A (summarized in USP Drug Information for the Health Care Professional). Of particular importance are the numerous drugs that inhibit or compete for CYP 3A, because they increase FK-506 blood levels and lead to toxicity (Prograf package insert, Fujisawa US, Rev 4/97, Rec 6/97). Also important are the drugs that induce P450 3A (e.g. Dexamethasone), because they decrease FK-506 blood levels and reduce efficacy. Because the major site of CYP 3A action on FK-506 is removed in the analogs provided by the present invention, those analogs are not as susceptible to drug interactions as the naturally occurring compounds.

Hyperglycemia, nephrotoxicity, and neurotoxicity are the most significant adverse effects resulting from the use of FK-506 and are believed to be similar for FK-520. Because these effects appear to occur primarily by the same mechanism as the immunosuppressive action (i.e. FKBP-calcineurin interaction), the intrinsic toxicity of the desmethoxy analogs may be similar to FK-506. However, toxicity of FK-506 is dose related and correlates with high blood levels of the drug (Prograf package insert, Fujisawa-US, Rev 4/97, Rec 6/97). Because the levels of the compounds provided by the present invention should be more controllable, the incidence of toxicity should be significantly decreased with the 13-desmethoxy analogs. Some reports show that certain FK-506 metabolites are more toxic than FK-506 itself, and this provides an additional reason to expect that a CYP 3A resistant analog can have lower toxicity and a higher therapeutic index.

Thus, the present invention provides novel compounds related in structure to FK-506 and FK-520 but with improved properties. The invention also provides methods for making these compounds by fermentation of recombinant host cells, as well as the recombinant host cells, the recombinant vectors in those host cells, and the recombinant proteins encoded by those vectors. The present invention also provides other valuable materials useful in the construction of these recombinant vectors that have many other important applications as well. In particular, the present invention provides the FK-520 PKS genes, as well as certain genes involved in the biosynthesis of FK-520 in recombinant form.

FK-520 is produced at relatively low levels in the naturally occurring cells, *Streptomyces hygroscopicus* var. *ascomyceticus*, in which it was first identified. Thus, another benefit provided by the recombinant FK-520 PKS and related genes of the present invention is the ability to produce FK-520 in greater quantities in the recombinant host cells provided by the invention. The invention also provides methods for making novel FK-520 analogs, in addition to the desmethoxy analogs described above, and derivatives in recombinant host cells of any origin.

The biosynthesis of FK-520 involves the action of several enzymes. The FK-520 PKS enzyme, which is composed of the *fkfA*, *fkfB*, *fkfC*, and *fkfP* gene products, synthesizes the core structure of the molecule. There is also a hydroxylation at C-9 mediated by the P450 hydroxylase that is the *fkfD* gene product and that is oxidized by the *fkfO* gene product to result in the formation of a keto group at C-9. There is also a methylation at C-31 that is mediated by an O-methyltransferase that is the *fkfM* gene product. There are also methylations at the C-13 and C-15 positions by a methyltransferase believed to be encoded

by the fkbG gene; this methyltransferase may act on the hydroxymalonyl CoA substrates prior to binding of the substrate to the AT domains of the PKS during polyketide synthesis. The present invention provides the genes encoding these enzymes in recombinant form. The invention also provides the genes encoding the enzymes involved in ethylmalonyl CoA and 2-hydroxymalonyl CoA biosynthesis in recombinant form. Moreover, the invention provides *Streptomyces hygroscopicus* var. *ascomyceticus* recombinant host cells lacking one or more of these genes that are useful in the production of useful compounds.

The cells are useful in production in a variety of ways. First, certain cells make a useful FK-520-related compound merely as a result of inactivation of one or more of the FK-520 biosynthesis genes. Thus, by inactivating the C-31 O-methyltransferase gene in *Streptomyces hygroscopicus* var. *ascomyceticus*, one creates a host cell that makes a desmethyl (at C-31) derivative of FK-520. Second, other cells of the invention are unable to make FK-520 or FK-520 related compounds due to an inactivation of one or more of the PKS genes. These cells are useful in the production of other polyketides produced by PKS enzymes that are encoded on recombinant expression vectors and introduced into the host cell.

Moreover, if only one PKS gene is inactivated, the ability to produce FK-520 or an FK-520 derivative compound is restored by introduction of a recombinant expression vector that contains the functional gene in a modified or unmodified form. The introduced gene produces a gene product that, together with the other endogenous and functional gene products, produces the desired compound. This methodology enables one to produce FK-520 derivative compounds without requiring that all of the genes for the PKS enzyme be present on one or more expression vectors. Additional applications and benefits of such cells and methodology will be readily apparent to those of skill in the art after consideration of how the recombinant genes were isolated and employed in the construction of the compounds of the invention.

The FK-520 biosynthetic genes were isolated by the following procedure. Genomic DNA was isolated from *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) using the lysozyme/proteinase K protocol described in Genetic Manipulation of *Streptomyces* - A Laboratory Manual (Hopwood *et al.*, 1986). The average size of the DNA was estimated to be between 80 - 120 kb by electrophoresis on 0.3% agarose gels. A library was constructed in the SuperCos™ vector according to the manufacturer's instructions and with the reagents provided in the commercially available kit (Stratagene). Briefly, 100 µg of

genomic DNA was partially digested with 4 units of *Sau3A* I for 20 min. in a reaction volume of 1 mL. and the fragments were dephosphorylated and ligated to SuperCos vector arms. The ligated DNA was packaged and used to infect log-stage XL1-BlueMR cells. A library of about 10,000 independent cosmid clones was obtained.

5 Based on recently published sequence from the FK-506 cluster (Motamedi and Shafiee, 1998, *Eur. J. Biochem.* 256: 528), a probe for the *fkbO* gene was isolated from ATCC 14891 using PCR with degenerate primers. With this probe, a cosmid designated pKOS034-124 was isolated from the library. With probes made from the ends of cosmid pKOS034-124, an additional cosmid designated pKOS034-120 was isolated. These cosmids
10 (pKOS034-124 and pKOS034-120) were shown to contain DNA inserts that overlap with one another. Initial sequence data from these two cosmids generated sequences similar to sequences from the FK-506 and rapamycin clusters, indicating that the inserts were from the FK-520 PKS gene cluster. Two *EcoRI* fragments were subcloned from cosmids pKOS034-124 and pKOS034-120. These subclones were used to prepare shotgun libraries by partial
15 digestion with *Sau3A*I, gel purification of fragments between 1.5 kb and 3 kb in size, and ligation into the pLitmus28 vector (New England Biolabs). These libraries were sequenced using dye terminators on a Beckmann CEQ2000 capillary electrophoresis sequencer, according to the manufacturer's protocols.

To obtain cosmids containing sequence on the left and right sides of the sequenced
20 region described above, a new cosmid library of ATCC 14891 DNA was prepared essentially as described above. This new library was screened with a new *fkbM* probe isolated using DNA from ATCC 14891. A probe representing the *fkbP* gene at the end of cosmid pKOS034-124 was also used. Several additional cosmids to the right of the previously sequenced region were identified. Cosmids pKOS065-C31 and pKOS065-C3
25 were identified and then mapped with restriction enzymes. Initial sequences from these cosmids were consistent with the expected organization of the cluster in this region. More extensive sequencing showed that both cosmids contained in addition to the desired sequences, other sequences not contiguous to the desired sequences on the host cell chromosomal DNA. Probing of additional cosmid libraries identified two additional
30 cosmids, pKOS065-M27 and pKOS065-M21, that contained the desired sequences in a contiguous segment of chromosomal DNA. Cosmids pKOS034-124, pKOS034-120, pKOS065-M27, and pKOS065-M21 have been deposited with the American Type Culture Collection, Manassas, VA, USA. The complete nucleotide sequence of the coding

sequences of the genes that encode the proteins of the FK-520 PKS are shown below but can also be determined from the cosmids of the invention deposited with the ATCC using standard methodology.

Referring to Figures 1 and 3, the FK-520 PKS gene cluster is composed of four open reading frames designated *fk bB*, *fk bC*, *fk bA*, and *fk bP*. The *fk bB* open reading frame encodes the loading module and the first four extender modules of the PKS. The *fk bC* open reading frame encodes extender modules five and six of the PKS. The *fk bA* open reading frame encodes extender modules seven, eight, nine, and ten of the PKS. The *fk bP* open reading frame encodes the NRPS of the PKS. Each of these genes can be isolated from the cosmids of the invention described above. The DNA sequences of these genes are provided below preceded by the following table identifying the start and stop codons of the open reading frames of each gene and the modules and domains contained therein.

	<u>Nucleotides</u>	<u>Gene or Domain</u>
15	complement (412 - 1836)	<i>fk bW</i>
	complement (2020 - 3579)	<i>fk bV</i>
	complement (3969 - 4496)	<i>fk bR2</i>
	complement (4595 - 5488)	<i>fk bR1</i>
	5601 - 6818	<i>fk bE</i>
20	6808 - 8052	<i>fk bF</i>
	8156 - 8824	<i>fk bG</i>
	complement (9122 - 9883)	<i>fk bH</i>
	complement (9894 - 10994)	<i>fk bI</i>
	complement (10987 - 11247)	<i>fk bJ</i>
25	complement (11244 - 12092)	<i>fk bK</i>
	complement (12113 - 13150)	<i>fk bL</i>
	complement (13212 - 23988)	<i>fk bC</i>
	complement (23992 - 46573)	<i>fk bB</i>
	46754 - 47788	<i>fk bO</i>
30	47785 - 52272	<i>fk bP</i>
	52275 - 71465	<i>fk bA</i>
	71462 - 72628	<i>fk bD</i>
	72625 - 73407	<i>fk bM</i>
	complement (73460 - 76202)	<i>fk bN</i>
35	complement (76336 - 77080)	<i>fk bQ</i>
	complement (77076 - 77535)	<i>fk bS</i>
	complement (44974 - 46573)	CoA ligase of loading domain
	complement (43777 - 44629)	ER of loading domain
	complement (43144 - 43660)	ACP of loading domain
40	complement (41842 - 43093)	KS of extender module 1 (KS1)
	complement (40609 - 41842)	AT1
	complement (39442 - 40609)	DH1
	complement (38677 - 39307)	KR1
	complement (38371 - 38581)	ACP1

SUBSTITUTE SHEET (RULE 26)

	complement (37145 - 38296)	KS2
	complement (35749 - 37144)	AT2
	complement (34606 - 35749)	DH2 (inactive)
	complement (33823 - 34480)	KR2
5	complement (33505 - 33715)	ACP2
	complement (32185 - 33439)	KS3
	complement (31018 - 32185)	AT3
	complement (29869 - 31018)	DH3 (inactive)
	complement (29092 - 29740)	KR3
10	complement (28750 - 28960)	ACP3
	complement (27430 - 28684)	KS4
	complement (26146 - 27430)	AT4
	complement (24997 - 26146)	DH4 (inactive)
	complement (24163 - 24373)	ACP4
15	complement (22653 - 23892)	KS5
	complement (21420 - 22653)	AT5
	complement (20241 - 21420)	DH5
	complement (19464 - 20097)	KR5
	complement (19116 - 19326)	ACP5
20	complement (17820 - 19053)	KS6
	complement (16587 - 17820)	AT6
	complement (15438 - 16587)	DH6
	complement (14517 - 15294)	ER6
	complement (13761 - 14394)	KR6
25	complement (13452 - 13662)	ACP6
	52362 - 53576	KS7
	53577 - 54716	AT7
	54717 - 55871	DH7
	56019 - 56819	ER7
30	56943 - 57575	KR7
	57710 - 57920	ACP7
	57990 - 59243	KS8
	59244 - 60398	AT8
	60399 - 61412	DH8 (inactive)
35	61548 - 62180	KR8
	62328 - 62537	ACP8
	62598 - 63854	KS9
	63855 - 65084	AT9
	65085 - 66254	DH9
40	66399 - 67175	ER9
	67299 - 67931	KR9
	68094 - 68303	ACP9
	68397 - 69653	KS10
	69654 - 70985	AT10
45	71064 - 71273	ACP10

1 GATCTCAGGC ATGAAGTCCT CCAGGCGAGG CGCCGAGGTG GTGAACACCT CGCCCTCTCT
 61 TGTACGGACC ACTTCAGTCA GCGGCGATTG CGGAACCAAG TCATCCGGAA TAAAGGGGCG
 121 TTACAAGATC CTCACATTGC GCGACCGCCA GCATACGCTG AGTTGCCTCA GAGGCAAACC
 181 GAAAGGGCGC GGGCGGTCCG CACCAGGGCG GAGTACCGCA CGAGAGTGGC GCACCTGCGC

241 ACCGTACACT CTCTCCCCCG CCGGCGGGAT GCGCGGCGTG ACACGGTTGG GGTCTCTCTG
 301 ACCGTGAACA CCGCGCGCGT GTGGCGTCGG GGACACCGCC TGGCATCGGC CCGGTGACCG
 361 TACGGGGAGG GCGTACGGCG GCGGTGGCTC GTGCTACCG CCGCCGGGCG GTCATCCCTG
 421 GAGACGSCAC TCGGCGAGCA GGGACGCGTG GTGCTACCT GCGGGCGGGA CGACCGTCTG
 481 GTTCCCGGGG GGGCGGTGGC CCGTGGTGAG CCACTCTCTC AGGGCGGTGA AGCTGAGCG
 541 GTGACACGGC AGCAAAGGCC GGAGTCGGTC GGGGAAAGTG TCGACGAGGG CGTCCGGTGTG
 601 CCGCGCGCTC TCGATGCGST AGTAGCGGTA CCGGCTGCA GCGCGCTGCC GGACATACCG
 661 GCGTACACGT CCGAGCCCGG GCGGCAGGCA GCAGCACTTC GAGAGTGCCT GGATGGTGTAT
 721 CAGCGGCTTC CCGATACGAC CCGTCAACGC GATGCTTTC ACGGCCGCGT GGACGCCCGA
 781 GGAGCGGGTG GCGTAGTCGT AGTCGGCATC GCAGCCCGGG ACCGTCCCCG GGGCGCAATA
 841 CCGTGTGCCG GCTTCTTTCT CCCCATCGAA GCGGGGGTCG AACTCCTCGC GGTAGACGCG
 901 CTGCGTCAGA TCCAGTAGA CCTCGTGGTG GTACGGCCAC AAGAACTCGG AGTCGGCCGG
 961 GAGCCCGGCG CCGAGCAGCG CCTCGCGCGC CTGGCCGGCT GCGGGGCGCG CTGCCGCGTA
 1021 GGTGGGGTAG TCGCGCAGGG CCGCCGGCAG GAAGGTGAAG AGGTTGGGAC CCTCCGCGCG
 1081 CCACAGGGTG CCTTCCCAGT CCACTCCTCC GTGCTACAGC TCGGGATGGT TCTCCAGCTG
 1141 CCAGCGCAGC AGGTAGCCCG CGTTGGACAT CCGGTGACC AGGGTGCCTG CGAGCGGCCG
 1201 GTGGTAGCGC TGGGCGACCG ACGCGCGGGC GGCCCGGGTC AGCTGGGTGA GCGGGGTGT
 1261 CCACTCGGCG ACGGCGTCGC CCGCCCGGGA GCCATCACGG TAGAACGGTG GGCCGGTGT
 1321 GCGCTTGTCT GTGGCGGCGT AGGCGTAACC GCGGGCGAGC ACCAGTCCG CGATGGCCCG
 1381 GTGCTTGGCG TACTGCTCGG GGTACCGGG GGTGCCGGCC ACGACCAGGC CACCGTTCCA
 1441 GCGGTCGGGC AGCCGGATGA CGAACTGGGC GTGCTGGTTC CACCCGTGGT TGGTGTGGT
 1501 GGTGGAGGTG TCGGGGAAGT AGCCGTGCTG CTGGATCCCG GGCACCTCCG TGGGAGTGGC
 1561 CAGTCTCTTG GCGGTGAGCC CTGCCAGTC CCGCGGCTCG GTGTGGCCCG TGGCCGCGCT
 1621 TCCCGCGGTG GTACGCTCGT CCAGCGAGTC GCGCTGCTGA CGTGGCCCG CCGGGACAGC
 1681 CAGCTGGGAC AGACGGGCGC AGTGACCGTC CCGGGGCTCG GGAGCAGGCC GCGCGGCGG
 1741 CCGTGAGGGG AGCAGGACCG CCACTGCGGC CAGGCTGAGA GCGCCGAGGC CCGTGCCTCT
 1801 TCTCGGGGCC CGTCCGACAC CGAGGGGCGA AACCATGGAG AGCCTCCAGA CGTGCGGATG
 1861 GATGACGGAC TGGAGGCTAG GTGCGGCACG GTGGAGACGA ACATGGGTGC GCGCGCATG
 1921 ACTGAGGCC CTCAGAGGTG GGCCGCGGCC ATGACGGGCG CCGGACCGCG GCGCTCCCG
 1981 GCGCGGTGCC GCGCCGCCCA CCGGTTCGCG GTCCCGGGT CAGGGACAGG TGTGCTTCG
 2041 GACGGTGAAG TAGCCGGTCG CCACTCTTTT CAAGGTGGTC GTGACGAAGG TGTGTACAG
 2101 GCGCATGTTT TGGCCGGAGC CTTTGGCGTA GGTGTAACCG GCGCTCGTCG TGGCGCGGCC
 2161 GCGCTGGACG TGAGCGTAGT TGCCGGCGGT CCAGCAGACG GCGGTGGCAC CCGTCTCTG
 2221 CCGGTTGACC GCGCCCGAGA GCGGTCCGCG CTTGCCGTCC GCGTCCCGGG CCGCGACCGC
 2281 GTAGGTGTGC GATGTGCCCC CECTCAGGCC GGTGTCCGTG TACGACGTCT TGGCGGACGT
 2341 GGTGATCTGG GCACCGTCGC GGTGGACCGG GTAGTCGGTG GCGCCGTCTG CCGGTTTCCA
 2401 GGTGAGGCTG ATGGTGGTGT CCGTGGCGCC GGTGGCGGCC AGGCGGACG GAGCGGGCAG
 2461 CGAACCAGGG TCGGAGGCGG ATCCGCTCAG CCGGAAGAAC TGCGTAGTCC AGTACGGGA
 2521 ACAGATCGAG TCCAGGAAGT AGGCGGCGCC GGTGCTGCCG CACTGCTGTG CTCCGGTGGC
 2581 GGGATCGACC GGGGTGCCCT GCGCGATGCC CCGCACCCCG TTACCTCCA CCGCCACCGA
 2641 TCGCTCCGCG GCCAGGTACT CCTCGTGCCG GGTGGAGTTC GGGCGGATCA CCGAGGTACG
 2701 GTCCGGCGTC TGGGACACGC CGTGCACAGC GGTCCACTGG TCGCGCAACT CGTGCGCGTT
 2761 GCGCGGCGCG ACGGTGGTGT CTTGTTCGCC GTGCCAGATG GCCACGCGCG GCCACGGGCC
 2821 CGACACGAG GGGTAGCCGT CACGACCCG CCGCCGCCAC TGGTCCCGCG TCAGGTCCGT
 2881 CCGGGGGTTC ATGCACAGGT ACGCGCTGCT GACGTGGTG GCACAGCGGA AGGGCAGGCC
 2941 GCGGACGACC GCGCCGGCCT GGAAGACGTC CCGATAGGTG GCGAGCATCA CCGACGTCAT
 3001 GGCACCGCCG GCGGACAGCC CCGTGATGTA GGTGCCCTGG GGGTCCGCGC CGTAGGCGGA
 3061 GACGGTGTGA GCGGCCATCT GCGGATCGA CCGGGCTTCG CCCTGGCCCC TCGGTTCTC
 3121 GCTGCTCTGG AACCAGTTGA AGCACCTGTT CCGGTTGTTT GACGACGTGG TCTCGGCGAA
 3181 CACGAGCAGG AAGCCATAGC GGTCCGCGAA TGAGAGCAGG CCGGAGTTGT CCGCGTAGCC
 3241 CTGGCGTCC TGGGTGCAAC CGTGACGGGC GAACACCACC GCGGCTCCG CCGGACGGA
 3301 CCGGGGCGCG TAGACGTACA GTTTCAGCCG GCGCGGGTTC GTGCCGAAGT CCGGACCTC
 3361 GGTGAGGTCC GCCTTGGTCA GACCGGGCTT GCGCAGGCC GCGCGGCGGT GGGCGGTCTC
 3421 CCGCGGGGCG AGCAGGGCCG CTCCGAGTAC GAGGGCCACG ACGGCCACGA GACGGGTGAG
 3481 CACCCCCCGC CGTCCCGGAC GCGACAACGA CCGGACCGGC GCGGAGGAGG AGAGGGGGA
 3541 CAGCGGGGTG AGGATTCCCC GGAACGGCGG CCGCTGCATG GCGGCTCCCT CGATGTCTG
 3601 GGCGGGACAC GGAGGGCTCC CTGACGTCGA TCACTGGGAG CGCCCCGGTG CCGGCGACCG
 3661 TCGGGGTGT TCAACCCGCA ACGGTATGCG CCGGAGCACC ACACCCGCA CCGGCGATG
 3721 TCGCGCCGGA CGGATTGTCT GCGCTTGGCG AATCTGATAC CCGGACCGCA CGAACGCCCG
 3781 ACCCGACACG GGTAGGGCGT CATGGTGTCC GACTCGGCG GTGCGCCTTG CCTGCCCTGG

SUBSTITUTE SHEET (RULE 26)

3841 ACGGACCGGG CGTCGGCGGA CCGGGCGTCG GCGGGCTGGG CGGTATGGCG GCGGAGGACG
3901 CCAGCCGCGT GGGGCGGGCG CGCCCAAGTG CAGTACGCGG ACCGTGGGCG GCGGGAGGGC
3961 CGGACCGGTC AGTGCAGTCC CGCGGCCCTG CCGGACCGCT CGTCCCAGAC GGGTTCCACC
5 4021 GCGGCGAACC GGGGTCCGTG TCCGCGGGCG TAGACCATCA GTGTCCGCTC GAGGGTGATC
4081 ACGATGACAC CGTCCTGGTT GTAGCCGATG GTGCGCACGC TGATGATGCC TACGTCAGGT
4141 CGGCTGGCGG ACTCCCGGGT GTTCAGGACC TCGGACTGCG AGTAGATGGT JTCGCCCTCG
4201 AAGACCGGGT TCGGCAGCCT GACCCGGTCC CAGCCGAGGT TGGCCATCAC ATGCTGGGAG
4261 ATGTGCGTGA CGCTCTGCCC GGTGACCAGG GCGAGGGTGA AGGTGGAGTC CACCGACGGC
10 4321 TTGCCCCAGG TGGTGCCCGC CGAGTAGTGG CCGTCAAGT GCAGCGGCGC GGTGTTCTCC
4381 GTCAGGAGCG TGAGCCAGGA GTTGTGGTTC TCCAGGACCG TCGGGCCAG GGGGTGGCGG
4441 TACACGTGCG CGGTGGTGAAG GTCTCGAAG TAGCGGCCCT GCCAGCCCTC GACCACAGCG
4501 GTGCGGGTGG CGTCTGGTTC CCGGTTCTCA GTCGTCATGG CGCTCATTCT GGGAGTCCC
4561 CGGTCCGCTG TGAATGCCG AACCTTCACC GGGCTCATAC GTGCGGCGCA TGAGCCCTGG
4621 ACCGTACGTA GTCGTAGAAC CTGCCACCA CTGGCGCGG TGGTCTCTCC GCGAGTGTGA
15 4681 CCACGCCGAC CGTGCGCCGC GCTGCGGGT CGTCGAGCGG CACGGCGACG GCGTGGTCA
4741 CCGGCCCGGA CCGGCTGCCG GTGAGGGGGG CGACGGCCAC ACCGAGGCCG CCGCGCACG
4801 GGGCCCGCAG CGTGCTCAGC TCGGTGCTCT CCAGGACGAC CCGCGCACG AATCCGGCCG
4861 CCGCGCACAG CCGGTCCGTG ATCTGGCGCA GTCCGAAGAC CCGTCCAGT GCCACGAACG
4921 CCTCATCGGC CAGCTCCGCG GTCCGCACCC GCGGCGTCT GGCCAGCCGG TGTCCGGGTG
20 4981 GGAGGAGCAG GCACAGTGCC TCGTCCCGCA GTGGTGTCCA CTCCACATCG TCCCGGGCGG
5041 GTCGTGGGCT GGTGAGCCC AGGTCCAGCC TGCTGTTGCG GACGTGTCG ACCACGGCGT
5101 CCGCGCGCTC GCGCGCGAGT TCGAAGGTGG TCGCGGGAGC CAGCGGGCGG TACCGGGCGA
5161 GGAGGTCCGG CACCGGCCAG GTGCCGTAGG AGTGCAGGAA ACCCAGTGCC ACGGTGCCG
5221 TGTCGGGGTC GATCAGGGCG GTGATGCGCT GCTCGGCGCC GGAGACCTCA CTGATCGCGC
25 5281 GCAGGGCGTG GCGCGGGAAG ACCTCGCCGT ACTTGTGAG CCGGAGCCGG TTCTGGTGCC
5341 GGTGCAACAG CCGCACGCCC ACTCGTCGCT CCAGCCGCCG GATGGCCCTG GACAGGGTCC
5401 GCTGGAGAT GTTGAGCCGT TCCGCGGTGA TCGTCACGTG CTCGTGCTCG GCCAAGGCCG
5461 TGAACCACTG CAACTCCCGT ATCTCCATGC AGGACTATA CGTACCGGGC ATGGTCTCTG
30 5521 CGAGGTTTCG TCATTTTACA GCGGCCGGGC GCGGGCCAC AGTGAGTCCT CACCAACAG
5581 GACCCCATGG GAGGGACCCC ATGTCCGAGC CGCATCCTCG CCCTGAACAG GAACGCCCG
5641 CCGGGCCCTT GTCCGGTCTG CTCGTGGTTT CTTTGGAGCA GGCCGTGCGC GCTCCGTTCC
5701 CCACCCGCCA CCTGGCGGAC CTGGGCGCCC GTGTATCAA GATCGAACGC CCGGCAGCG
5761 GCGACCTCGC CCGCGGCTAC GACCGCACGG TCGGTGGCAT GTCCAGCCAC TTCGTCTGGC
35 5821 TGAACCGGGG GAAGGAGAGC GTCCAGTCCG ATGTGCGCTC GCGGAGGGG AACCGGCACC
5881 TGCACGCCTT GGTGGACCGG GCCGATGTC TGGTGCAGAA TCTGGACCC GCGCGCGCG
5941 GCGGCCTGGC ATCGGCCACC AGGTCTCTCG GCGGAGCCAC CGAGGCTGAT CACCTGCGGA
6001 CATATCCGGC TACGGCAGTA CCGGCTGCTA CCGCGGACCG CAAGGCGTAC GACCTCTTGG
6061 TCCAGTGCGA AGCGGGGCTG GTCTCCATCA CCGGCACCCC CGAGACCCCG TCCAAGGTGG
40 6121 GCCTCTCCAT CGCGGACATC TGTGCGGGGA TGTACGCGTA CTCGGGCATC CTCACGGCCC
6181 TGCTGAAGCG GGCCCGCACC GGCCGGGGCT CGCAGTTGGA GGTCTCGATG CTCGAAGCCC
6241 TCGGTGAATG GATGGGATAC GCCAGTACT ACACGCGCTA CGGCGGCACC GCTCCGGCCC
6301 GCGCCGGCGC CAGCCACGCG ACGATCGCCC CTACGCGCCC GTTACCACG GTTCCGGGC
6361 AGACGATCAA TCTCGGGCTC CAGAACGAGC GGGAGTGGGC TTCTTTCTGC GGTGTCGTGC
45 6421 TACAACGCCC CGGTCTCTGC GACGACCCCG GCTTTTCCGG CAACGCCGAC CCGGTGGCGC
6481 ACCGCACCGA GCTCGACGCC CTGGTGAGCG AGGTGACGGG CACGCTCACC GCGGAGGAAC
6541 TGGTGGCGCG GCTGGAGGAG GCGTCGATCG CCTACGCACG CCAGCGCACC GTGCGGGAGT
6601 TCAGCGAACA CCCCCAATG CGTGACCGTG GACGCTGGGC TCGTTTCGAC AGCCCGGTGG
6661 GTGCGCTGGA GGGCCTGAT CCCCCGGTCA CCTTCCACGG CGAGCACCCG CCGCGGCTGG
50 6721 GCGGGGTCCC GGAGCTGGGC GAGCATACCG AGTCCGTCTT GCGGTGGCTG GCGCGGCC
6781 ACAGCGCCGA CCGCGAAGAG GCCGGCCATG CCGAATGAAC TCACCGGAGT CCTGATCCTG
6841 GCGCGCGTGT TCCTGCTCGC CGGCGTACGG GGGCTGAACA TGGGCTGTCT CCGGCTGGTC
6901 GCCACCTTTC TGCTCGGGGT GGTGCGACTC GACCGAACGC CGSACGAGGT GCTGGCGGGT
6961 TTCCCCGCGA GCATGTTCTT GGTGCTGGTC GCGGTCACGT TCCTCTTCGG GATCGCCCCG
7021 GTCAACGGCA CCGTGGACTG GCTGGTACGT GTCGCGGTGC GGGCGGTGGG GGCCCGGGTG
55 7081 GGAGCCGTCC CTTGGGTCTT CTTCGGCTG CCGGCACTGC TCTGCGCGAC AGGCGCGGCC
7141 TCGCCCGCGG CCGTGGCGAT CTGGGCGCG ATCAGCGTCG CGTTCGCGGT CAGGACCCGC
7201 ATCGATCCGC TGTACGCGG ACTGATGGCG GTGAACGGGG CCGCAGCCGG CAGTTTCCGC
7261 CCTCCGCGA TCCTGGGCGG CATCGTCCAC TCGGCGCTGG AGAAGAACCA TCTGCCCGTC
7321 AGCGGCGGGC TGCTCTTCGC AGGCACCTTC GCCTTCAACC TGGCGGTGCG CCGGTTGTCA
60 7381 TGGCTCGTCC TCGGGCGCAG GCGCCTCGAA CCACATGACC TGGACGAGGA CACCGATCCC

	7441	ACCGGAAGGGG	ACCCGGCTTC	CCGCCCCGGC	GCGGAACACG	TGATGACGCT	GACCGCGATG
	7501	GCCGCGCTGG	TGCTGGGAAC	CACGGTCCTC	TCCCTGGACA	CCGGCTTCCT	GGCCCTCACC
	7561	TTGGCGGCGT	TGCTGGCGCT	GCTCTTCCCG	CGCACCTCCC	AGCAGGCCAC	CAAGGAGATG
5	7621	GCCTGGCCCG	TGGTGTGCTG	GGTATGCGGG	ATCGTGACCT	ACGTGCGCCT	GCTCCAGGAG
	7661	CTGGGCAATG	TGGACTCCCT	GGGGAAGATG	ATCGCGGCGA	TCGGCACCCG	GCTGCTGGCC
	7741	GCCCTGCTGA	TCTGCTACGT	GGGCGGTGTC	GTCTCGGCCT	TCGCCTCGAC	CACCGGGATC
	7801	CTCGGTGCCC	TGATGCCGCT	GTCCGAGCCG	TTCCTGAAGT	CCGGTGCCAT	CGGGACGACC
	7861	GGCATGCTGA	TGGCCCTGGC	GGCCGCGGCG	ACCGTGGTGG	ACGCGAGTCC	CTTCTCCACC
10	7921	AATGGTGCTC	TGGTGGTGCG	CAACGCTCCC	GAGCGGCTGC	GGCCCGGCGT	GTACCAGGGG
	7981	TTGCTGTGCT	GGGCGCCCGG	GGTGTGCGCA	CTGGCTCCCG	CGGCCGCGTG	GGCGGCGCTC
	8041	GTGGTGGCGT	GAGCGCAGCG	GAGCGGGGAT	CCCTTGAGAG	CCGTTTCCCG	TGCTGTGTCG
	8101	CTGACGTAGC	GTCAAGTCCA	CGTGCCGGGC	GGGCACTACG	CCTAGCATGT	CGGGCATGGC
	8161	TAATCAGATA	ACCCTGTCCG	ACACGCTGCT	CGCTTACGTA	CGGAAGGTGT	CCCTGCGCGA
15	8221	TSACGAGGTG	CTGAGCCGGC	TGCGCGCGCA	GACGGCCGAG	CTGCCGGGCG	GTGGCGTACT
	8281	GCCGGTGCAG	GCCGAGGAGG	GACAGTTCCT	CGAGTTCCTG	GTGCGGTTGA	CCGGCGCGCG
	8341	TACAGTGTCT	GAGATCGGGA	CGTACACCGG	CTACAGCACG	CTCTGCCTGG	CCCGCGGATT
	8401	GGCGCCCGGG	GGCCGTGTGG	TGACGTGCGA	TGTCATGCCG	AAGTGGCCCC	AGGTGGGCGA
	8461	GCGGTACTGG	GAGGAGGCCG	GGGTGTGCGA	CCGGATCGAC	GTCCGGATCG	CGGACGCCCG
20	8521	GACCGTCCTC	ACCGGGCTGC	TCGACGAGGC	GGGCGCGGGG	CCGGAGTCGT	TCGACATGGT
	8581	GTTCATCGAC	GCCGACAAGG	CCGGCTACCC	CGCCTACTAC	GAGGCGGCGC	TGCCGCTGGT
	8641	ACGCCGCGGC	GGGCTGATCG	TCGTGACAAA	CACGCTGTTC	TTCGGCCGGG	TGGCCGACGA
	8701	AGCGGTGCAG	GACCCGGACA	CGGTGCGGGT	ACGCGAACTC	AACGCGGCAC	TGCGCGACGA
	8761	CGACCGGGTG	GACCTGGCGA	TGCTGACGAC	GGCCGACGGC	GTCACCTCC	TGCGGAAACG
25	8821	GTGACCGGGG	CGATGTGCGG	GCGGCTCAGC	GTACAGCTCG	TCGGCGCGGG	CCTCGCGGAG
	8881	GGCTCCAGAT	GCAGGCGTTC	GACGCCGGCG	GCGGAAGCGC	CCGCCACCTC	GGACACGCAG
	8941	GGGCACTCGG	AGTCCGCGAA	GCCCGCGAAC	CGGTAGGCGA	TCTCCATCAT	GCGGTTGCGG
	9001	TCCGTACGCC	GGAAGTCCGC	CACCAGGTGC	GCCCCCGCGC	GGGCGCCCTG	GTCCGTGAGC
	9061	CAGTTCAGGA	TCGTGCGACC	GGCACCAGAC	GACACGACCC	GGCAGGACGT	GGCGAGCAGT
30	9121	TTCAGGTGCC	ACGTGCGACG	CTTCTTCTCC	AGCAGGATGA	TGCCGACGGC	GCCGTGCGGG
	9181	CCGAGCGGCT	CGCCCATGGT	GACGACGAGG	ACCTCATGGG	CGGGATCGGT	GAGCACGCGC
	9241	GCAAGTCCGC	GTCCGAGTAG	TGCACGCCGG	TCGCGTTTAT	CTGGCTGGTC	CGCAGCGTCA
	9301	GTTCTCTGAC	GCGGCTGAGT	TCCTCTCTCC	CCGCGGGTGC	GATCGTTCATG	GAGAGGTCGA
	9361	GCGAGCGCAG	GAAGTCCTCG	TCGGGACCGG	AGTACGCCTC	CCGGGCGCTG	TCGCGCGCGA
35	9421	AACCCGCGTG	GTACATCAGG	CGGCGCCGAC	GCGAGTCGAC	CGTGGACACC	GGCGGGCTGA
	9481	ACTCCGGCAG	CGACAGGAGC	GTGGCCGCTT	GCTCGGCCGG	GTAGCACCAG	ACCTCGGGCA
	9541	GGTGGAAACG	CACCTCGGCA	CGCTCGGCGG	GCTGGTCTGC	GATGAACGCG	ATCGTGGTCC
	9601	GTGCGAAGTT	CAGTCCGTG	GCGATCTCGG	GGACGGACTG	CGACTTCGGC	CCCCATCCGA
	9661	TGCGGGCCAG	CACGAAGTAC	TCCGCCACAC	CGAGGCGTTC	CAGACGAGGT	CACGCGAGGT
40	9721	CGTGGTCTGT	CTTGCTCGCC	ACCGCCTGGA	GGATGCCGCG	GTGCTCGAGC	GTGGTGATCA
	9781	CCTCGCGGAT	CTCGTGGGTG	AGGACCACCT	CGTCTCTCTC	CAGCACGCTG	CCCCGCCACA
	9841	AGGTGTTGTC	CAGGTCCGAG	ACCAGACACT	TGACAATGGT	CATGGCTGTC	CTCTCAAGCC
	9901	GGGAGCGCCA	GCGCGTGCTG	GGCCAGCATC	ACCCGGCACA	TCTCGCTGCT	GCCCTCGATG
	9961	ATCTCCATGA	GCTTGGCGTC	GCGGTACGCC	CGTTCCAGCA	CGTGTCCCTC	TCTCGCGCCT
45	10021	GCCGACGCGA	GCACCTGTGC	GCGGTCGCGG	GCCCGGGCGG	CGGCTCGTTC	GGCGGCGAGC
	10081	TGCTTGGCCA	GGATCGTTCG	GGGCACCATC	TCGGGCGAGC	CCTCGTCCCA	GTGGTCTGCT
	10141	GCGTACTCGC	ACACGCGGGC	CGCGATCTGC	TCCGCGGTCC	ACAGGTGCGC	GATGTGCCCG
	10201	GCGACGAGTT	GGTGGTCGCC	GAGCGGCGCG	CCGAACTGCT	CCCGGGTCCG	GGCGTGGGCC
	10261	ACCGCGGCGG	TGCGGCAGGC	CCGCAAGGATC	CCGACGCGAG	CCGAGGCGAC	CGACTTGCGC
50	10321	CCGTAGGCGA	GTGACGCCGC	GACCAGCATC	GGCAGTGACG	CGCCGGAGCC	GGCCAGGACC
	10381	GCGCCGCGCG	GCACACGCAC	CTGGTCCAGG	TGCAGATCGG	CGTGGCCGGC	GGCGCGGCAG
	10441	CCGGACGCGT	TGCGGACGCG	CTCGACGCGT	ACGCCGGGGG	TGTCGGCGGG	CACGACCACC
	10501	ACCGCACCGG	AACCATCCTC	CTGGAGACCG	AAGACGACCA	GGTGGTCCGC	GTAGGCGGGC
	10561	GCACTCGTCC	AGACCTTGTC	GCCGTGACGC	ACAGCGGTGT	CCCCGTGAG	CCGAACCCGC
	10621	GTCCGCATCG	CCGACAGATC	GCTGCCCGCC	TGCCGCTCAC	TGAAGCCGAC	GGCCGCGAGT
55	10681	TTCCCGCTGG	TCAGTCTCTT	CAGGAAGGTC	GCCCCGCTGAC	CGGCGTCCGC	GAGCCGCTGC
	10741	ACGGTCCACG	CGGCCATGCC	CTGCGACGTC	ATGACACTGC	GCAGCGAATC	GCAGAGGCTG
	10801	CCGACGTGTC	CGGTGAACTC	GCCGTTCTCC	CGGCTGCCGA	GTCCAGAGAC	CCGCTGCTCC
	10861	CCGCGCACTT	CCGCGCAGAG	CAGGCTCTCG	GCGCGGAGCC	GGACCGAGCC	GTCGCGCGGC
	10921	AGTTCGCGCG	ACGTGTCCCA	CTCGGCGGCC	CGGTACCGCA	CAAGGTCCGT	CAGCAGCGCG
60	10981	TCACGCTCAG	GCATCGACGG	CCCGCAGCCG	GTGGACGAGT	GCGACCATGG	ACTCGACGGT

SUBSTITUTE SHEET (RULE 26)

11041 ACGGAAGTTC GCGAGCTGGA GGTCCGGGGC GGGGATCGTG ACGTCGAACG TCTTCTCCAG
 11101 GTACACGACC AGTTCCATCG CGAACAGCGA CGTGAGGCGG CCTCCGCGA ACAGGTCCGG
 11161 GTCCACGGGC CAGTCCGACC TGGTCTTCGT CTTGAGGAAC GCGACCAACG CGTGCGCGAC
 11221 GGGGTCTGTC TTGAGGGGTG CGGTCTATGA AACACCTTCT CGTATTCGTA GAAGCCCCCG
 5 11281 CCGGTCTTCC GGCCGTGGTG TCCCTCGCGG ACCTTGCCCA GCAGCAGGTC ACAGGGGGCG
 11341 CTGGGCTCGT CGCCGGTGCG TTTGTGCAGC ACCACAGCG CGTCGACGAG GTTCTCGATG
 11401 CCGATCAGGT CCGCGGTGCG CAGCGGGCCG GTCGGATGGC CGAGGCACCC CGTCATGAGC
 11461 GCGTCGACGT CCTCGACGGA CGCGGTGGCC TCCTGCACGA TCCGCGCCCG GTCGTTGATC
 11521 ATCGGGTGGA GCAGCCGGCT CGTGAAGGAG CGGGGCGCGT CCGGACGAC CATCGGCTTG
 10 11581 CCGCCGAGCG CCGCGAGCAG GTCCCGGGCG GCGGCCATGG CCTTCTCACC GGTCCGGGT
 11641 CCGCGGATCA CCTCGACCGT CGGGATCAGG TACGACGGT TCATGAAGTG CGTCCCGAGC
 11701 AGGTCTTCGG GCCGGGCCAC GGAGTGGGCG AGTTCGTCAA CCGGGATCGA CGACGTGTTT
 11761 GTGATGACCG GGATACCGGG CGCCGCTGCG GAGACCGTGG CGAGTACCTC CGCCTTGACC
 11821 TCGCGTCTCT CGACGACGGC CTCGATCACC GCGGTGGCCG TACCGATCGC GGGCAGCGCG
 15 11881 GACGTGGCCG TCCGCAGCAC ACCGGGCTCG GCCTCGCGG GCCCGGCCAC GAGTGTGTC
 11941 GTCCGCAAGT CCGGTGGCGAT CCGCGCCCGC CCGCCCGTAA GGATCTCCTC GGACGTGTCG
 12001 ACGAGTGTCA CCGGGACGCC GTGGCGCAGC GCGAGCGTGG TGATGCCGGT GCCCATCACT
 12061 CCGCGCCCGA GCACGATCAG CTGGTGGTCC ACGCTGTTTC CTCCTCCCGG GTCCACCATG
 12121 GCAGCGAGTA CGGGTGCGAG ACGTCTTCGG GGGTCGACCC GATCGCGTCC TTGCGGCCGA
 20 12181 GGCCGAGTTC GTGCGCGAAG CCGAGCAGCA CGTCGAACGC GATGTGGTCG GCGAACGCGC
 12241 TGCCCGTCGA GTCGAGGACG CTCAGGCTGT CCGGTGGTCC GCGCGCGGTG TCCGGTGCCG
 12301 CGCACAGGGC CGCCAGCGAC GGGCCGAGCT CGCGGTCCCG CAGTTGTCTG TACTCGCCCT
 12361 CGGCGCGGGC CTGCCCCGGA TGGTCGACGC AGATGAACGC GTCGTCGAGC AGGCTCTTCG
 25 12421 GCAGTTCGGT CTTGCCCGGC TCGTCGGCGC CGATGGCGTT CACATGCAGG TCGCGCAGCC
 12481 GCGGCTCGGC GGGCAGCACC GGCCCTTTGC CCGAGGGCAC CGAGGTGACG GTGGACAGGA
 12541 CATCCGCGGC GCGCGCGGCC TCCGCCCGAT CGGTCACTT GACCGGCAGT CCGAGGAACG
 12601 CGATGCGGTC CGCGAACGAC GCGCGCTGSC CCGGGTCCGT GTCGCTGACC AGGATCCGCT
 12661 GATGGGCGAG GACCTGTCTG AGCGCTGCG CCTGGGTCAC CGCTGTGCG CCCGCGCCGA
 30 12721 TCAGCGTGAG CGTGGCGCTG TCGGACCGGG CACGACGCC GCGCGGACCG CGCGCGACCG
 12781 CGCCGGTCCG CATCGCGGTG ATCAGCCTG CGTCGGCGAG GCGCGTCAGA CTGCCGTGT
 12841 CGTCGTCGAG GCGCGACATC GTGCCGACGA TCGTCGGCAG CCGGAAGCGC GGATAGTTGT
 12901 GCGGACTGTA CGAAACCGTC TTCATGGTCA CGCCGACACC GGGGACCCGG TACGGCATGA
 12961 ACTCGATGAC GCCGGGAATG TCGCCGCCGC GGACGAATCC GGTACGCGGC GCGCCTCGG
 35 13021 CGAACTCGCC GCGGCCGAGC GCGGCGAACC CGTCGTGAG CTCGCTGATC AGCCGGTCCA
 13081 TCATCAGTC GCGGCCGATC ACGGAGAGAA TCCGCTTGAT GTACGTTGTG CGCAGGACCC
 13141 TGGTCTGCAT GTGTACCTC CTTTCTGTG CGGAGCTGT CTTGGTGGT GCGCTCGGG
 13201 CGGCTTCCGT TCTCATCGCA GCTCCCTGTC GATGAGGTG AAAATCTCGT CCGCGGTGCG
 13261 GTCCGCGGAC AGCACGCCGG CCGGCTGTGT CCGGCGGGTC TCCCGCCGCC AGCGTTGAG
 40 13321 CAGGGCGTCC AGCCGGGTTT CGATCGCGTC CGCCTGGCG GCGCCCGGGT CGACACCGGC
 13381 AACGAGTGCT TCCAGCCGGT CGAGCTGCGC GAGCACCACG GTCACCGGGT CGTCCGGGGA
 13441 CAGCAGTTCA CCGATGCGGT CGGCGAGTGC GCGCGGCGAC GGGTAGTCGA AGACGAGCGT
 13501 GCGCGACAGT CGCAGACCGG TGCCTCGTT GAGGCCGTTG CGCAGCTGCA CCGCGATGAG
 13561 CGAGTCCACA CCGAGTTCCC GGAACGCCGC GTCCTCCGG ATGTCCTCCG GTGCGGCGTC
 45 13621 GCGCAGGACG GCGCTGCCT TCTGCCGAC GAGGGCGAGC AGGTGCGTGG GCGTTCTTG
 13681 CTCGTTGCGG GCGCTCCGGC GGGCCGACGG CTTGGGCGCG CCACGCGACA GCGGGAGGTC
 13741 CCGCGGCGAG TCGCCCGCCA CGGCGACGAC ACTGCCCGTT CCGGTGTGGA CCGCGGCGTC
 13801 GTACATGCGC ATGCCCTGTT CCGCGGTGAG CGCGCTCGCC CCACCCTTGC GCATACGGCG
 13861 CCGGTCCGGC TCGGTCAAGT CCGCGGTGAG GCCACTCGCC TGGTCCCACA GCGCCACGCG
 50 13921 GATCGACAGC CTTGGCAGCC CTTGTGCAAG CGGCTGTTG GCGAGCGCGT CGAGGAACGC
 13981 GTTCCCGGCC GCGTAGTTGC CCTGACCGGG GGTGCCCAGC ACACCGGCCG CCGACGAGTA
 14041 GACGACGAAT GCGGCGAGGT CCGGTGTGCG GGTGAGCCGG TGCAGGTGCC AGGCGGCGTC
 14101 GGCCTTGGGT TTGAGGACGG TGTGATGCG GTCCGGGGTG AGGTTGTGCA GCAGGGCGTC
 14161 GTCGAGGGTT CCGGCGGTGT GGAAGACGCG GGTGAGGGGT TGAGGGATGT GGGCGAGGGT
 55 14221 GGTGGCGAGT TGGTGGGGGT CCGCGACGTC GCAGGGGAGG TGGGTGCCGG GGGTGGTGT
 14281 GGGGGGTGGG GTGCGGGAGA GGAGGTAGGT GTGGGGGTGG TTCAGGTGGC GGGCGAGGAT
 14341 CCGCGGAGG GTGCGGGAGC CGCCGCTGAT GACGACGGCC CCTCGGGGT CCAGCGGCCG
 14401 CCGGACCGTG AGGACGATCT TGCCGCTGTG CTCGCCCGCG CTCATGGTCG CCAGCGCCTC
 14461 GCGGACCTGC CGCATGTCGT GCACCGTCAC CGGCGCGGG TGCAGCACAC CGCGCGCGAA
 14521 CAGGCCGAGC AGCTCCGCGA TGATCTCCTT GAGCCGGTGC GGCCCGCGCT CCATCAGGTC
 60 14581 GAACGGTCCG TGGACGGCGT GCCGGATGTC CTTCTTCCCC ATCTCGATGA ACCGGCCACC

SUBSTITUTE SHEET (RULE 26)

	14641	CGGCGCGAGC	AGGCCGACGG	ACGCGTCGAG	GAGTTCACCG	GTGAGCGAGT	TGAGCACGAC
	14701	GTCGACCGGC	GGGAACGCGT	CGGCGAACGC	GGTGCTGCGG	GAATCGGCCA	GATGCGCTCC
	14761	GTCCAGGTCC	ACCAGATGGC	GCTTCGCGGC	GCTGGTGGTC	GCGTACACCT	CGCGCGCCAC
5	14821	GTCCCGCGCG	ATCTGCCGGG	CGGCGGAACC	GACACCGCCG	GTGGCGCGCT	GGATCAAGAA
	14881	CTTCTCGCCG	GGGCGCAGCC	CGGCGAGGTC	GACCAGGCCG	TACCACGCGG	TCGCGAACGC
	14941	GGTCATCACG	GACGCCGCCT	GCGGGAACGT	CCAGCCGTCC	GGCATCCGGC	CGAGCATCCG
	15001	GTGGTCCGGC	ATGACCGTGG	GGCCGAAGCC	GGTGCCGACG	AGGCCGAAGA	CGCGGTGCGC
	15061	CGGTGCGAGA	CCGGAGACGT	CGGCGCCGGT	CTCCAGGACG	ATGCCCGCGG	CCTCGCCGCC
10	15121	GAGCACGCCG	TGACCGGGGT	AGGTGCGCAG	CGCGATCAGC	ACATCGCGCA	AGTTGAGGCC
	15181	CGCGCGACGC	ACACCGATCC	GGACCTCGGC	CGGGGCGAGG	GGGCGCCGGG	GCTCCGCCGA
	15241	GTGGGCCCGG	GTGAGGCCGT	CGAGGGTGCC	CGTCCGCGCC	GGCCGGATCA	GCCACGTGTC
	15301	GCTGTCCCGC	ACGGTGAGCG	GCTCCGGCAC	CGGGGTGAGG	CGGGCCGCCT	CGAACCAGCC
	15361	GCCGCGCAGC	CGCAGACGCG	GCTCGCCGAG	TGCGACGCGG	ATGCGCTGCT	GCTCGGGGGC
15	15421	GAGCGTGACG	CCGGACTCGG	TCTCGACGTG	GACGAACCGG	CCGGGCTGCT	CGGCCCTGGG
	15481	GCGCGCGCAG	AGTCCGGCCG	CGGCGCCGGT	GGCGAGGCCG	GCGGTGGTGT	GCACGAGCAG
	15541	ATCCCCGCGG	GAGCCGGTCA	GGGCGCGTAG	CAGCCGGGTG	GTGAGCGCAC	CGGTCTCGGC
	15601	CACCGGGTCG	TCGCCATCAG	CGGCAGGCAA	CGTGATGACG	TCCACGTGCG	TCGCGGGGAC
	15661	ATCCGTGGGT	GCGGCGACCT	CGATCCAGGT	GAGACGCATC	AGGCCGGTGC	CGACGGGTGG
20	15721	GGACAGCGGG	CGGGTGCGGA	CCGTCCGGAT	CTCGGCGACG	AGTTGGCCGG	CGGAGTCGGC
	15781	GACGCGCAGA	CTCAGCTCGT	CGCCGTACAG	AGTGATCAGC	GCTCGGAGCA	TGGCCGAGCC
	15841	CGTGGCGACG	AACCGGGCCC	CCTTCCAGGC	GAACGGCAGA	CCCGCAGCGC	TGTCGTCCGG
	15901	CSTGGTGAGG	GCGACGGCGT	GCGGCGCCGC	GTGAGCAGC	GCCGGATGCA	CACCGAAACC
	15961	GTCCGCGCTG	GCGGCCTGCT	CGTCCGGGAG	CGCCACCTCG	GCATACACGG	TGTACCATC
25	16021	ACGCCAGGCA	GCCCCGAACC	CCTGGAACGC	CGACCCGTAC	TCATAACCGG	TATCCCCGAG
	16081	TTCGTATAG	AACCCCGAGA	CGTCGACGGC	CACGGCCGTG	ACCGGCGGCC	ACTGCGAGAA
	16141	CGGCTCCACA	CCGACAACAC	CGGGGGTGTC	GGGGGTGTG	GGGGTCAGGG	TGCCGTGGC
	16201	GTGCCGGGTC	CAGCTGCCCC	TGCCCTCGGT	ACGCGCGTGG	ACGGTCAACG	GCCGCCGTCC
	16261	GGCTTCATCA	GCCCCCTTCCA	CGGTCAACCA	CACATCCACC	GCTGCGGTCA	CCGGCACCCAC
30	16321	AAGGGGGGAT	TCGATGACCA	GCTCGTCCAC	TATCCCGCAA	CCGTCTCGT	CACCGGCCCC
	16381	GATGACCAGC	TCCACAAACG	CCGTACCCGG	CAGCAGGACC	GTGCCCCGCA	CCCGCTGATC
	16441	AGCCAGCCAG	GGGTGAGTGC	GCAATGAGAT	CCGGCCAGTG	AGAACAACAC	CACCATCGTC
	16501	GGCGGGCAGC	GCTGTGACAG	CGGCCAGCAT	CGGATGCGCC	GCACCCGTCA	ACCCCGCCGC
	16561	CGACAGATCG	GTGGCACCGG	CCGCCTCCAG	CCAGTACCGC	CTGTGCTCGA	ACGCGTACGT
35	16621	GGSCAGATCC	AGCAGCCGTC	CCGGCACCGG	TTCGACCACC	GTGTCCCAGT	CCACTGCCCT
	16681	GCCCAGGGTC	CACGCCGTGC	CCAACGCCGT	CAGCCACCGC	TCCCAGCCGC	CGTCACCGGT
	16741	CCGCAACGAC	GCCACCGTGT	GGAACCTGCTC	CATCGCCGGC	AGCAGCACCG	GATGGCGACT
	16801	GCACTCCACG	AACACCGACC	CATCCAGCTC	CGCCACCGCC	GCGTCCAACG	CCACCGGACG
	16861	ACGCAGATTC	CGGTACCACT	ACCCCTCATC	CACCGGCTCC	GTCACCCAGG	CGCTGTCCAC
40	16921	GGTCGACCAC	CACGCCACCG	ACGCGGCCTT	CCCTGCCACC	CCCTCCAGTA	CCTTGGCCAG
	16981	TTCATCCTCG	ATGGCTTCCA	CGTGGGGCGT	GTGGGAGGCG	TAGTCGACCG	CGATACGACG
	17041	CACCCGCACG	CCTTCGGCCT	CATACCGCGC	CACCACCTCC	TCCACCGCCG	ACGGGTCCCC
	17101	CGCCACCAAC	GTCGAAGCCG	GGCGGTTACG	CGCCGCGATC	CACACACCTT	CGACGAGACC
	17161	GACCTCACCG	GCCGGCAACG	CCACCGAAGC	CATCGCTCCC	CGCCCGGCCA	GTCGCGCCGC
45	17221	GATGACCTGA	CTGCGCAATG	CCACCACGCG	GGCGGCGTCC	TCGAGGCTGA	GGGCTCCGGC
	17281	CACGCACGCC	GCCGCGATCT	CGCCCTGGGA	GTGTCCGATC	ACCGCGTCCG	GCACGACCCC
	17341	ATGCGCCTGC	CACAGCGCGG	CCAGGCTCAC	CGCGACCGCC	CAGCTGGCCG	GCTGGACCAC
	17401	CTCCACCCGC	TCCGCCACAT	CCGGCCGCGC	CAACATCTCC	CGCACATCCC	AGCCCGTGTG
	17461	CGGCAGCAAC	GCCTGAGCGC	ACTCCTCCAT	ACGCGCGGGC	AACACCGCGG	AGTGGGGCAT
50	17521	GAGTTCCACG	CCCATGCCGA	CCCCTGGGCG	GCCCTGGCCG	GGGAAGACGA	ACACCGTACG
	17581	CGGCTGGTCC	ACCGCCACAC	CCGTACCCCG	GGCATCGCCC	AGCAGCACCG	CACGGTGACC
	17641	GAAGACAGCA	CGCTCCCGCA	CCAACCCCTG	CGCGACCGCG	GCCACATCCA	CACCAACCCC
	17701	GCGCAGATAC	CCCTCCAGCC	GCTCCACCTG	CCCCCGCAGA	CTCACCTCAC	CACGAGCCGA
	17761	CACCGGCAAC	GGCACCAACC	CGTCAACAAC	CGACTCCCCA	CGCGACGGCC	CAGGAACACC
55	17821	CTCAAGGATC	ACGTGCGCGT	TCGTACCGCT	CACCCCGAAC	GACGACACAC	CCGCATGCGG
	17881	TGCCCCGATC	GACTCGGGCC	ACGGCCTCGC	CTCGGTGAGC	AGCTCCACCG	CACCGGCGCA
	17941	CCAGTCCACA	TGCGACGACG	GCTCGTCCAC	ATGCAGCGTC	TTCGGCGCGA	TCCCGTACCG
	18001	CATCGCCATG	ACCATCTTGA	TCACACCGGC	GACACCCGCC	GCCGCCTGCG	CATGACCGAT
	18061	GTTGCACTTC	AACGAACCCA	GCAGCAGCGG	AACCTCACGC	TCCTGCCCGT	ACGTGCGCAG
60	18121	AATGGCCTGC	GCCTCGATGG	GATCGCCCGC	CGTCGTCCCC	GTCCCGTGCG	CCTCCACCAC
	18181	GTCCACATCG	GCGGCGCGCA	GTCCGGCGTT	CACCAACGCC	TGCTGSATGA	CACGCTGCTG

SUBSTITUTE SHEET (RULE 26)

18241 GGACGGGCGG TTGGGGGCGG ACAGCCCGTT GGAGGCACCG TCCTGGTTCA CCGCCGACCC
18301 GCGGACGACC GCGAGAACGG TGTGTCCGTT GCGCTCGGCG TCGGAGAGCC GCTCCAGCAC
18361 AAGAACGCGG GCGCCCTCCG CCCAGCCGGT GCGGTTGGCG GCGTCCGCGA ACAGCGGGCA
5 18421 GCGGCCGCTG GGGGAGAGTC CGCCCTGCTG CTGGAATTCC ACGAACCCGG TCGGGGTCG
18481 CATGACGGTG ACACCGCCGA CCAGCGCCAG CGAGCACTCC CCGTGGCGCA GTGCGTGCCC
18541 GGCCTGGTGC AGCGCGACCA GCGACGACGA GCACGCCGTG TCCACCGTGA ACGCCGGTCC
18601 CTGGAGCCCA TAGAAGTACG AGATCCGGCC GGTGAGCACG CTGGGCTGCA TGCCGATCGA
18661 GCCGAACCCG TCCAGGTCCG CGCCGACGCG GTACCCGTAC GAGAAGGCGC CCATGAACAC
18721 GCGGGTGTGG CTGCCGCGCA GTGTGCCCGG CACGATGCCC GCGCTCTCGA CGCCTCCCA
10 18781 TGTCGTTTCC AGCAGGATCC GCTGCTGGGG GTCCATGGCC CGTGCCCTAC GGGGGCTGAT
18841 GCCGAAGAAC GCGGCATCGA AGCCGGCGGC GTCGGAGAGG AAGCCGCCGC GGTCCGTGTC
18901 CGATCCGCGG GTGAGGCCGG ACGGGTCCCA GCCACGGTCG GCCGGGAAGC CCGTGACCGC
18961 GTCGCCGCCA CTGTCCACCA TGCGCCACAG GTCGTCCGGC GAGGTGACGC CGCCCGGCAG
15 19021 TCGCAGGCGG ATGCCACGA TGGCCAGCGG TTCGTACCGG GTCGCGCGCG CTGTGGGAAC
19081 AGCGACCGGT GCGGCACCA CCGCCAGAGC CTCGTCCAAC CGCGAGCGA TGGCCCGCGG
19141 CGTCGGGTAG TCGAAGACAA GCGTGGCGGG CAGTCGGACA CCGGTGCGCG CGCGAGTCG
19201 GTTCCGCACT TCGACGGCGG TCAGCGAGTC GATACCCAGT TCCTTGAAGG CCGCGTCCGC
19261 GGACACGTCC GCGGCGTCCG CGTGCCGAG CACCGCCGCC GCGTTGTGCG GGACAGTGC
19321 CAGCAGCGCG GTGTCCCGCT CAGCGCCGGA CATGGTGCCG AGCCGGTCCG CGAGCGGAAC
20 19381 GCGGGTGGCC GCCGCCGGGC GCGATACGGC GCGGCGCAGA TCGGCGAAAA GCGGCGATGT
19441 GTGCGCGGTG AGGTCCATCG TGGCCGCCAC GCGGAACGCG GTGCCGGTTC CCGCCCGCGG
19501 TTCCAGCAGG CGCATGCCCC CACCCGCGCA CATGGGGCGG AAACCGCGAC AGCGGACAG
19561 GGTGCGGTTG GTGCCGCTCA TGCTGCCGGT GAGTCCGCTG TCATCGGCCC AGAGGCCCA
25 19621 GGCCAGCGAC AGCGCGGGCA GTCCTTCGGC ATGGCGCAGC GTCGCGAGTC CGTCGAGGAA
19681 CCCGTTCCGC GCCGAGTAGT TGCCCTGGCC GCGGCCGCC ATGATGCCCG CGACGGACGA
19741 GTAGAGGACG AACGAGCGCA GGTCCGCGTC CCGGGTCAGC TCGTGCAAGT GCCAGGCGCC
19801 GTCGGCTTTG GGGCGCAGTG TGGTGGCGAG CCGCTCCGGG GTGAGTGCCG TGGTCACGCC
19861 GTCGTCGAGC ACGGCTGCCG TGTGGAAGAC CGCCGTGAGC GGCCTGCCG CGGCGCGAG
30 19921 CGCGGCGGCG AGCTGGTCCC GGTGCGGAC GTACAGCGG ATGTGGACAC CGGAGTGTG
19981 CGCCGGCGGT TCGCTGCGCG ACAGCAACAG GAGGTGGCGG GCGCCATGCT CGGCGACGAG
20041 ATGCCGGGCG AGGAGACCTG CCAGCACACC CGAGCCGCGG GTGATGACCA CCGTGCCGTC
20101 CCGGTCGAGC AGCGGTTCCG GCGTTTCCGC GCGGCGCGTG CCGGTGAACC GCGGCGCTTC
20161 GTACCGGCCG TCGGTGACGC GGACGTACGG CTCGGCCAGT GTCGTGGCGG CGGCCAGCGC
35 20221 CTCGATGGGG GTGTGCGTGC CGGTCTCCAC CAGCACGAAC CCGCCCGGGT GCTCGGCTG
20281 GCGCGACCGG ACGAGGCGCG GCAGCGCTCC TCCGACCGGT CCGCGCTCGA TCCGGACGAC
20341 GAGGGTGCTC TCCGCAAGGC CGTCTCGGC GATCACCCGG TGCAGTCCG CGGAGCGAA
20401 CTCGGTGAGC CGGTACGTCT CGTCGAGGAC ATCCGCGCCC GGTTCGGGA GCGCGGAGAC
20461 GATGTGGACC GCGTCCGCGG GACCGGGCCC GGGAGTGGGC AGCTCGGTCC AGGAGAGGCC
40 20521 GTACAAGGAG TTCCGTACGA CCGCGGCGTC GCGGTCGACG TTCACCGGTC GCGCGTCCG
20581 CGCGGCGAGC GTCACCACCG GTTGGCCGAC CCGGTCCGTC GCATGCACGG CAGCGCCGTC
20641 CCGGCCCTGA GTGATCGTGA CGCGCAGCGT GGTGGCCCCG GTCGTGTGGA ACCGCAACGC
20701 GCTCCAGAG AACCGCAGC GCACCTCCGC TTCCTGTTCC CGGAGCAGG GCAGGACGGT
20761 GACGTGCAAG GCCGCGTCGA ACAGCGCCGG TTGGACGCCA TAGTGCGGCG TGTCTCCG
45 20821 CTGTTCCCGG GCGATCTCCA CCTCGGCGTA CAGGGTTTCG CCGTCGCGCC AGCGGTGCG
20881 CAGTCCCTGG AACGCTGGGC CGTAGCTGTA GCCGGTCTCG GCCAGCCGCT CGTAGAACGC
20941 GCTACGTCG ACGCGTCGCG CGCCCGGCGG CCGCCACGCG GCGGCGGGGA CCGCCGCGAC
21001 GCTTCCGGCC CGGCCGAGGG TGCCGCTGGC GTGCCGGGTC CAGCTGTCCG TGCCCTCGGT
21061 ACGCGGCTGG ACGGTCACTC GCCCGCGTCC GGCCTCATCG GCGCCTTCGA CCGTCAACGA
50 21121 CACATCCACC GCGCCGCTCA CCGGACCAAC GAGCGGGGTC TCGATGACCA GTTCATCCAC
21181 CACCCCGCFA CCGGTCTCGT CACCGGCCCG GATGACCAGC TCCACAAACG CCGTACCCGG
21241 CAGCAGAACG GTGCCCCGCA CCGCGTGATC AGCCAGCCAG GGATGCGTAC GCAACGAGAT
21301 CCGGCCASTG AGAACAACAC CACCACCGTC GTCGGCGGGC AGTGCTGTGA CCGCGGCCAG
21361 CATCGGATGC GCGGCCCCCG TCAGCCCGGC CCGGACAGA TCGGTGGCAC CCGCCCGCTC
21421 CAGCCAGTAC CGCCTGTGCT CGAACGCGTA GGTGGGCGA TCGAGCAGCC GTCCCGGCAC
55 21481 CGGTCGAGC ACGGTGTCC AGTCCACTGC CGTGCCAGG GTCCACGCTC GCGCAACGC
21541 CGTCAGCCAC CGTCCGAGC CGCCGTCACC GGTCCGCAAC GACGCCACCG TGTGAGCTG
21601 TTCCATCGCC GCGAGCAGCA CCGGATGGGC GCTGCACTCC ACGAACCGG ACCCTCCAG
21661 CTCCGCCACC GCGCGCTCCA GCGCGACGG GCGACGCGAG TTCCGCTACC AGTAGCCCTC
21721 ATCCACCGGC TCGGTCACCC AGGCGCTGTC CACCGTGGAC CACCAAGCCA CCGACCCGGT
60 21781 CCGCGCGGAA ATCCCTCCA GTACCTCGGC CAACTCGTCC TCGATGGCTT CCACGTGGGG

21841 CGTGTGGGAG GCGTAGTCGA CCGCGATAAG GCGCACTCGC ACGCCTTCGG CCTCGTACCG
 21901 CGTCACCACT TCTTCCACCG CGGACGGGTC CCCC GCCACC ACAGTCGAAG ACGGGCCGTT
 21961 ACGCGCCGCG ATCCACACGC CCTCGACCAG GTCCACCTCA CCGGCCGGCA ACGCCACCGA
 22021 AGCCATCGCC CCGCGCCCGG CCAGCCGCCC GGCGATCACC TGGCTGCGCA AGGCCACCA
 5 22081 GCGGGCGGCG TCCTCAAGGC TGAGGGCTCC GGCCACACAC GCGCGCGCGA TCTCGCCCTG
 22141 GGAGTGTCCG ACCACCGCGT CCGGCACGAC CCCATGCGCC TGCCACAGCG CGGCCAGGCT
 22201 CACCGCGACC GCCCAGCTGG CCGGCTGGAC CACCTCCACC CGCTCCGCCA CATCCGGCCG
 22261 CGCCAACATC TCCCGCACAT CCCAGCCCGT GTGCGGCAAC AACGCCCGCG CACACTCCTC
 22321 CATACGAGCC GCGAACACCG CATCAACTCC ACACCCATGC CCACCCACTG
 10 22381 AGCACCCCTG CCGGGAAAGA CGAACACCGT ACGCGGCTGA TCCACGCCCA CACCCATCAC
 22441 CCGGGCATCC CCAACAACA CCGCACGGTG ACCGAAGACA GCACGCTCAC GCACCAACCC
 22501 CTGCGCGACC GCGGCCACAT CCACACCACC CCGCGCAGA TACCCCTCCA GCGGCTCCAC
 22561 CTGCCCCCGC AGACTCACCT CACTCCGAGC CGACACCGGC AACGGCACCA ACCCATCGAC
 22621 AGCCGACTCC CCACGCGACG GCCCGGGAAC ACCCTCAAGG ATCACGTGCG CGTTCGTACC
 15 22681 GCTCACCCCG AAAGCGGAGA CACCGGCCCG GCGCGGACGT CCGCGCTCGG GCCACGCCCG
 22741 CGCTCGGTG AGCAGTTCCA CCGCGCCCTC GGTCCAGTCC ACATGCGACG ACGGCTCGTC
 22801 CACATGCAGC GTCTTCGGCG CGATGCCATA CCGCATCGCC ATGACCATCT TGATGACACC
 22861 GCGACACCCC GCAGCCGCCT GCGCATGACC GATGTTGAC TCAACGAAC CCAGCAGCAG
 22921 CGGAACCTCA CGCTCCTGCC CGTACGTGCG CAGAATCGCG TGCGCCTCGA TGGGATCGCC
 20 22981 CAGCGTCGTC CCGTCCCGT GCGCCTCCAC CACGTCCACG TCGGCGGGGG CGAGCCCCGC
 23041 CTTGTGGAGG GCCTGGCGGA TGACGCGCTG CTGGGAGGGG CCGTTGGGTG CGGAGATGCC
 23101 GTTGGAGGCG CCGTCTGGT TGACGGCGGA GGAGCGGACG ACCGCGAGGA CCGTGTGTCC
 23161 GTTCCGCTCG GCGTCGGAGA GCTTTTCGAC GACGAGGACG CCGGCCCTC CGCGAAACC
 23221 GGTGCCGTCC GCGCGCTCAG CGAACGCCCT GCACCGTCCG TCCGCGCGCA CGCCGCCCTG
 25 23281 CCGGGAGAAC TCCACGAAGG TCTGTGGTGA TGCCATCACT GTGACACCAC CGACCAGCGC
 23341 CAGCGAGCAC TCCCGGTCC GCAGCGCCTG CCGGCGCTGG TGCAGCGCGA CCAGCGACGA
 23401 CGAACACGCC GTGTGACCG TGACC3CCGG ACCCTCCATG CCGAAGAAGT ACGACAGCCG
 23461 TCCGGCGAGC ACCGCGGGCT GTGTGCTGTA GGCGCCGAAT CCGCCAGGT CCGCGCCCGT
 23521 GCGTAGCCG TAGTAGAAGC CGCCGACGAA GACGCCGGTG TCGTGCCTGC GCAGGGTGTG
 30 23581 CCGCACGATG CCGCGTGTG CCGCGCCTC CCAGGCGATT TCGAGGAGGA TCCGCTGTGT
 23641 CCGGTGAGT GCGGTGGCCT CGCGCGGACT GATGCCGAAG AACCGCGCAT CGAAGTCGGC
 23701 GCGCCCCGCG AGTGCGCCGG CCGCGCCGGT GGCGGACTCG GCGGCGGCGT GCAGCGCGGC
 23761 CACGTCCAG CCGCGGTGCG TGGGGAAGTC GCCGATCGCG TCGCGGCCGT CCGCGACGAG
 23821 CTGCCACAGC TCTTCCGGTG AGGTGACGCC GCCCGGCGAGT CCGCAGGCCA TGCCGACGAC
 35 23881 GCGAGCGGC TCGTTCGCCG CCGCGCGCAG CCGGGTGTTC TCCCGCGGGA GCTGCGCGTT
 23941 GTCCTTGACC GACGTCCGCA GCGCCTCGAT CAGGTGCTTC TCGGCCATCG CCTCATCCCT
 24001 TCAGCACGTG CCGCATGAGC CGCTCTGCGT CCATGTCTGC GAACAGTTCG TCGTCCGGCT
 24061 CCGCGGTCTG GGTGCTCGCG GGTGCTGTG CCGGTGGTTC ACCGCCGTCC GGGGTCCCGT
 40 24121 TGTCGTCCGG GGTCCCGTTG ACGTCCGGGG CCAGGAGGGT CAGCAGATGA CCGGTGAGCG
 24181 CGCCGGCGGC GGGATAGTCG AAGACGAGCG TGGCCGGCAG CGGAATGCCG AGGGCCTCGG
 24241 AGAGCCGGTT GCGCAGGCCG AGCGCGGTGA GCGAGTCGAC CCCGAGGTCC TTGAACGCCG
 24301 TGGTGGCCGT GACCGCCGCC GCGTCCGTGT GGCCAGCAG GGTGGCGGCG GTGTGCGGGA
 24361 CGACGCCGAG CAGCACCTGT TCCCGTTCTT TGTGGGGCAG GTCCGGCAGG CGTTCAGCA
 45 24421 GGGAGCCGCC GTCGGTCCGCG GAGCGCCGGG TGGGGCGCTG GATCGGTGCG CACAGCGGTG
 24481 ACGGGTCGCC GGGCCCGGGT GGGGCGGTG CCACGACCAC GGCTTCCCG GTGGCGCACG
 24541 CCGCGTCGAG GAGGTCCGTC AGCCGGTCCG CCGCGCGGGT GAACGCCACG GCCGGCAGGC
 24601 CTTGTGCCCC GCGCAGGTG GCGAGGGCCT GGAGCGGTCC GGCCGCTCG CCGGACGGAA
 24661 CCGCGAGAAC GAACGCGGTC AGGTGAGGT CCGGGTTCAG GCGGTGACGT TCCAGGCCG
 50 24721 ACTCGGCGGT CCGTCCCGG TGGACGACCG CGGTACCGG GGTTCGCGG ACTGTGCCCG
 24781 GTCGTACCG GATCACTTCG GCGCGGTG CCGCGAGGTG TCCGGCAGT TCCCGGAAC
 24841 CGCCCGCGAG GAGGACGGTG TCGCCGTACG AGGCCGCGGC CGTGGTGGGC GCGGCGGGGA
 24901 CGAGGCGGGG CGCTTCGAGG CGCCCGTCCG CCAGGCGCAG GTGCGGTTG TCGAGGCGGG
 24961 AGAGGGCGGC CCGCGCGCGG GGGGTGACCG TGTCGCTGCT CTCACGAGC ACGAGCCGGC
 55 25021 CCGGTTCCCG GGTGTCGAGC AGTGCGGCGA CCGCACCGGC GACGGGCCCG GCCTCGGCGG
 25081 ACACCACCAG CGTGGCGCCG GCGGTCTCTG GGTGCTCCAG TGCGGTACGG ACCTCGTCGG
 25141 GACCGGATAC CCGGACGACG ATGACGTCG CCGTGGCGTC GTCCCGAGG TCGGTGTACC
 25201 GCGGGCCGT GGTGCGGGT GCGCGCGGG CCGGACGCG GGTCCAGTG CGGTGAACA
 25261 CCGGCACGTC CCGTCCCGG CCGGTCTGCG GGGGGGCGG GGTGATGAGC GAGCCATCT
 25321 GAGCCACCGG CCGTCCAGT TCGTGGCGCA GGTGCACGCG GGCGCCGCC TCGCCCTCGC
 60 25381 CGTGACGAA GGTGACGCG AGTTTCGTGG CCGCGTGGT GTGGACACGG ACGCCGGTGA

SUBSTITUTE SHEET (RULE 26)

	25441	ACCGGAACGG	CAACCGTACC	CCCGCGTTCT	CGGCGGCGGC	GCCGATGCTG	CCCGCTTGCA
	25501	CGCGCGTGAC	GAGCAGCGCC	GGGTGCAGTG	TGTAGCGGGC	GGCGTCCCTG	GCGAGGGCGC
	25561	CGTCGAGGGC	GACTTCGGCG	CAGACGGTGT	CTCCGTGGCT	CCACGCGGCG	GACATGCCCG
5	25621	GGAACTCGGG	GGCGAACTCG	TATCCCGCGT	CGTCGAGTCG	CTGGTAGAAG	GCCGCGACCT
	25681	CGACCGGTTT	CGCGTGCTCG	GGCGGCCAGG	GCCCCGGCGT	GGTGGCCGGT	TCGGTGGTGG
	25741	CGATGCCGGC	GAAGCCGGAG	GCGTGCGGGG	TCCATGTCCG	GTCGCCGTCC	GTCCGGGCGT
	25801	GGACGCGCAC	GGCACGGCGT	CCGGTGTCTG	CGGGCGCGGC	GACGGTCAAC	CGCACCTGGA
	25861	CGGCGCCGGT	GGCGGGCAGG	ACCAGCGGTG	TCTCGACGAC	CAGTTCGTCT	AGCAGGTCCG
	25921	AGCCTGCCTC	GTCGGCGCCG	CGTCCGGCCA	ATTCCAGGAA	GGCGGGTCCG	GGCAGCAGTA
10	25981	CGCGCCGCTC	GACGGAGTGA	CCGGCCAGCC	ATGGGTGGGT	GGCCAGCGAG	AACCGGCCGG
	26041	TGAGCAGCAC	CTCGTCCGAG	TCGGGGAGCG	CCACCGACGC	GGCGAGCAGC	GGGTGGTCCA
	26101	CGCGCTCGAG	TCCGAGGCCG	GAAGCGTCCG	TGCCGGCCGC	GGTCTCGATC	CAGTAGCGCT
	26161	CATGGTGGAA	GGCGTATGTG	GGCAGGTCTG	GTGCCGTCTG	CGTCCGCGGG	ACGACCGCCG
	26221	CCAGTCCGAC	GGGCACGCCG	GTTGTGTGCG	CCTCGGCCAG	CGCGGTGAGC	AGCCGGTGGG
15	26281	CTCCCCCGCC	GCGGCGGAGC	GTGGCGACGG	TCCGCGCGTC	GATCGCGGGC	AGCAGCACGG
	26341	CGTGCGCGCT	GACCTCGACG	AACACGGTGT	CACCCGGCTC	GCGGGCAGCG	GTACCGGCCG
	26401	TGCGGAAGCC	TACGGGGTGG	CGCATGTTGC	GGAACCACTA	CTCGTCTCTG	AGCGCGCGCT
	26461	CGATCCAGCG	TTCGTCCGGC	GTGGAGAACC	ACGGGATCTC	GGGCGTGCGC	GAGGTGGTGT
	26521	CCCGGACGAT	CCGCTGGAGT	TCGTCTGACA	GCGGGTCGAC	GAACGGGGTG	TGGGTCCGGC
20	26581	AGTCGACGGC	GATGCGGGCG	ACCCAGACGC	CGCGGGCCTC	GTAGTCGGCG	ATCAGCGTTT
	26641	CGACGCGCTC	CGGGCGCCCG	GCGACGGTCT	TGGTGGTGGC	GCCGTTGCGG	CCCGCGACCC
	26701	AGACGCGCTC	GATCCGGGCG	GCATCCGCGT	CGACGTCCGC	GGCCGGGAGC	CGCACCGAGC
	26761	CCATCGCGCC	GCGTCCGGCG	AGTTCGCGCA	GGAGCAGGAG	AACGCTGCGC	AGCGCGACGA
	26821	GGCGGGCACC	GTCTCCAGG	GTGAGCGCTC	CGGCGACACA	GGCCGCGGCG	ATCTCGCCCT
25	26881	GCGAGTGTCC	GATGACGGCG	TCCGGGCGTA	CGCCCGCGGC	CTCCACACAG	GCGGCCAGCG
	26941	ACACCATGAC	GGCCAGCAG	ACGGGGTGCA	CGACGTGAC	GCGGCGGGTC	ACCTCCGGGT
	27001	CGTCGAGCAT	GGCGATGGGG	TCCCAGCCCC	TGTGCGGGAT	CAGCGCGTCT	GCGCATTTGG
	27061	GGTCCTTGGC	GGCGAACACC	GGGGAGGCCG	CCATCAGTTC	GACGCCCATG	CCGCGCCACT
30	27121	GCGGTCCCTT	TCCGGGGAAG	ACGAAGACGG	TGCGCGGCTC	GGTGAGCGCC	GTGCCCGTGA
	27181	CGACGTCTGC	GTCGAGCAGC	ACGGCGCGGT	GCGGGAACGT	CGTACGCGTC	GCGAGCAGGC
	27241	CCCGCGCGAT	GGCGCGCGGG	TCGTGGCCCG	GACGGGCGGC	GAGGTGCTCG	CGGAGTGGGC
	27301	GGACCTGGCC	GTCGAGGGCC	GTGGCGGTCC	GCGCCGAGAC	GGGCAGTGGT	GTGAGCGGCG
	27361	TGCGGATCAG	CGGCTCACCG	GGCTTCGAGG	CCGACGGCTC	CTCGGCGGGC	GGCTCCCCGG
35	27421	CCGGGTGGGC	TTCCAGCAGG	ACGTGGGCGT	TGGTGGCGCT	GACGCCGAAG	GAGGACACAC
	27481	CGCGCGCGCG	CGGGCGGTCT	GTCTCGGGCC	AGGGCCGGGC	ATCGGTGAGG	AGTTCGACCG
	27541	CGCGCGCGCT	CCAGTCGACG	TGCGAGGACG	GCGTGTCCAC	GTGACGGTGT	GCGGCGAGGG
	27601	TGCGCTGCGG	CATGGCGAGG	ACCATCTTGA	TGACACCGGC	GACACCCGCG	GCGGCTGAG
	27661	TGTGGCCGAT	GTTGGACTTC	AGCGAGCCCA	GCAGCACCGG	GGTGTGCGGC	CCCTGCCCGT
40	27721	AGGTGGCCAG	CACCGCCTGT	GCCTCGATGG	GATCGCCGAG	CCTGGTGGCG	GTGCCGTGCG
	27781	CCTCCACGGC	GTCCACGTCC	GCCGGGGTGA	GCCCGGCGGT	GGCCAGGGCC	TGCCGGATCA
	27841	CCCGCTCCTG	CGAGGGCCCC	TTCCGGCGCC	ACAACCCGTT	GGAAGCACCG	TCCTGGTTGA
	27901	CCGCCGAACC	CCGGACAACC	GCCAGCACAC	GGTGGCCGTT	GCGCTCGGCA	TCGGAGAGCC
	27961	TCTCGACGAT	CAGCACACCG	GACCCCTCGG	CGAAACCGGT	GCCGTACGCC	GCATCCGCGA
45	28021	ACGCTTTGCA	GCGCGCGTCT	GGCGCGAGAC	CCCGCTGCTG	GGAGAACTCG	ACGAAGCCGG
	28081	ACGGCGAGGC	CATCACCGTG	ACGCCGCCGA	CCAGGGCGAG	CGAGCATTCG	CCGGAGCGCA
	28141	GTGACTGCCC	GGCCTGGTGC	AGCGCCACCA	GCGACGACGA	ACACGCCGTG	TCGACCGTGA
	28201	CCCGCGGACC	CTCCAGACCG	TAGAAGTACG	ACAGCCGACC	GGACAGCACA	CTGGTCTGGG
	28261	TGCCGCTCGC	GCCGAAACCG	CCCAGGTCCG	TGCCGAGTCC	GTACCCGTCT	GAGAAGGCGC
50	28321	CCATGAACAC	GCCGGTGTCT	CTTCCGCGCA	GCGACTCCGG	GAGGATCCCG	GCGTGTTCCT
	28381	GCGCCTCCCA	CGAGGTCTCC	AGGACCAGAC	GCTGCTGCGG	GTCCATCGCC	AGCGCCTCAC
	28441	GCGGACTGAT	CCCGAAGAAC	GCGCGCTCGA	AGTCCGCCAC	CCCGGCGAGG	AAGCCACCAT
	28501	GACGCACGGT	CGACGTGCCC	GGATGATCCG	GATCGGGATC	GTACAGCCCC	TCCACGTCCC
	28561	AACCCACGGT	CGTCGGAAAC	GCGGTGATCC	CGTCACCAAC	CGACTCCAGC	AGCCGCCACA
55	28621	AGTCCCTCCG	CGACGCGACC	CCACCCGGCA	GCGGGCAGGC	CATCCCCACG	ATCGCCAAAC
	28681	GCTCGTCCCT	CCGGACGGCC	GCGGTCTGTC	TGCGGGTCCG	CGATGCCGTC	CGGCCGAGCA
	28741	GCGCGCGCGT	GAGCTTCGCC	GCGACGGCGC	GCGGGCTCCG	GAAGTCGAAG	ACCGCGGTGG
	28801	GCGCGAGCGG	TACGCCCCGT	GCGTGGGTGA	AGCGGTTGCG	CAGCGGATCG	CCCATGAGCG
	28861	AGTCGACGCC	GAGTTCCTTG	AACGTGCGCG	TGCGCTCGAC	CCGTGCGGCA	CCGTCTGGGC
60	28921	CGAGTACGGC	CGCGGTGCAC	TGCCGGACGA	CGGCGAGCAC	GTCTTTTTCG	GCGTCCGCGG
	28981	CGGAGAGCCG	GCGGATCCGG	TGCGCGAGGG	TGGTGGCGCC	GGCCGCCCGG	CGCCCGCGCT

SUBSTITUTE SHEET (RULE 26)

29041 CCCGGGCGCGG TGGCGGCGAGC AGGGGGCGAGC TGCCGAGGGCC GGCCGGGGTCG GCGGGCGACCA
29101 GCGCCGGGTCG CGAGGACCGC AACGCCGCGT CGAACAGCGT CAGTCCGCCCT TCGGGGGTCA
29161 GCGCCGTCAC GCGGTCGCGG CGCATGCGGG CGCCGGTSCC GACCGTCAGC CGGCTCTCCG
5 29221 GTTCCACACG GCGCCAGGCC ACGGACAACG CGGGCASTCC GGCTGCCCGG GCGTGTTCGG
29281 CCAGCGCSTC GAGGAACCGG TTCGCGGCCG CGTAGTTGCC CTGTCCGGGG CTSCCGAGCA
29341 CACCGGCGGC CGACGAGTAG AGGACGAACG CGGCCAGTTC CGTGTCTTGG GTGAGTTGGT
29401 GCAGGTGCCA CGCGGCGTCC ACCTTCGGGC GCAGCACCGT CTCGAGCCCG TCGGGGGTGA
29461 GCGCGGTGAG GACGCCGTCG TCGAGGACGG CGCGGGTGTG CACGACGGCC GTGAGCGGGT
10 29521 GCGCGGGTGC GATCCCCGCC AGTACGGAGG CGAGTTCTTC CCGGTCCGGC AGCTCCAGG
29581 CGATCCCGCT GACCTCGGCG CCGGGCACGT CGCTCCTCGT GCGGTGCGG GACAGCATCA
29641 GCAGCCGGCG CACGCCGTGG CGTTCGACGA GGTGGCCTCT GATGATGCCG GCCAGCCTCC
29701 CGGAGCCACC GGTGACGAGC ACGGTGCCGT CCGGTTCGAG CGCCGGAGCG TCACCCCGCG
29761 GGACCGCCCG GGCCAGACGG CCGGGCGTACA CCTGGCTCTC ACGCAGCACC ACCTGGGGCT
15 29821 CATCGAGCGC GGTGGCGGCT GCGAGCAGCG GCTCGGCTGG GTCCGGGGCG GCGTCGACGA
29881 GGACGATCCG GCGGGGGTGT TCGGCTGCG CGGTCCGAC CAGTCCGGCG GCGCGGGCGG
29941 ACGCGAGACC GGGCCCGGTG TGGACGGCCA GGACCGCGTC GGCGTACCGG TCGTCGGTGA
30001 GGAAGCGCTG CACGGCGGTC AGGACGCCGG CGCCCASTTC GCGGGTGTCT TCGAGCGGG
30061 CACCGCCCGC GCCGTGCGCG GGGAGGATCA CCACGTCCGG GACCGTCGGG TCGTCGAGGC
20 30121 GGCCGGTCTG CCGGTGCTGT GCGGGCAGCT CCGGGAGCTC GGCCAGCACC GGGCGCAGCA
30181 GGCCCGGAAC GGCTCCCGTG ATCGTCAGGG GGCGCTGCG CACGGCGCCG ATGGTGGCGA
30241 CCGGCCCGCG GGTCTCGTCC GCGAGGTGTA CGCCGTGAGC GGTGACGGCG ACGCGTACCG
30301 CCGTGGCGCG GGTGGCGTGG ACGCGGACGT CGTCGAACGC GTACGGAAGG TGGTCCCTCT
30361 CCGCGGCGAG GCGGAGTGGC GCGCCGAGCA GCGCGGGTGG CAGGCGGTAC CGTCCGGCTC
25 30421 CCGCGAGCTG TCCGTGCGCG AGGGCCACTT CCGCCAGAC GCGCTGCTCG TCGGCCGAGA
30481 CCGCGCGCGG GCGGGGCGAG GCGGGCCCGT CCGTGTACCC GGCTCGGGCC AGACGGTCTG
30541 CGATGTCTCT GGGGTCCACC GCGCGGGCGG TGGCGGGCGG CCACGTGAGC GGCATCTCC
30601 GCACGGCCCG GCGCGTCCCG GGTTCGGGGG CGAGGATTC GTGCGCGTGC TCGGTCCACT
30661 CCGCGCCCGC GTGCCGCGT TGCACGGTGA CCGCGCGCGG GCGGTCCGCC CCGGGCGCGC
30721 TCACCGTGAC CGAGAGCGCG AGCGACCGG ACCGCGSCAG CGTGAGGGGG GTGTCCACCG
30 30781 TGAACGTGTC GAGGGCGCGG CAGCCGCGTT CGTCGCGCGC CCGGATCCGG AGATCCAGGA
30841 GGGCCGCGGC GGGCAGCACC GCGAGGCGGT GCAGGGAGTG CGCCAGCGGA TCGGCGGCGT
30901 CGACCCGGCC GGTGAGCACC AGGTGCGCGG TGCCGCGCAG GGTGACCGCC GCGGTGAGCG
30961 CCGGGTGGCG GACCGGCGTC TGTCCGGCGG GGGCCGCTC GCGCGCGGTC TGGGTGCCGA
31021 GCCAGTAGCG GACCCGCTCG AACGGGTACG TCGGCGGGTG CGAGGCGCGT GCGGCGCGG
35 31081 GGTGATGAC CTTGCGCCAG TCGACCGTGA CGCCGTGCGT GTGCAGCCCG GCGAGCGCGG
31141 TCAGGGCGGA TCGGCTTCTG TCGTGGCGT GCAGCATCGG GATGCCGTG ACGAGTCGGG
31201 TCAGGCTCCG GTCCGGGCGG ATCTCCAGGA GCACCGCCCG CTCGTGCGCG GCGACCTGTT
31261 CCGCGAACCG GACGGTGTCT CCGACCTGTC GTACCCAGTA CTCCGGCGTG GTGCAGGCGG
40 31321 CCGCGCGCGC CATCGGGATC CTCGGCTCGT GGTACGTGAG GCTCTCCGCG ACCTTGCGGA
31381 ACTCCTCGAG CATCGGCTCC ATCCGCGCCG AGTGAACGC GTGGCTGGTC CCGAGCGGGG
31441 TGAAGCGGCC GAGCCGGGCG GCGACGTGCA GCACCGCTC CTCGTACCCG GAGAGCACGA
31501 TCGACGCGGG CCGTTGACC GCGGCGATCT CCACGCCGTC CCGCAGCAGC GGCAGCGCGT
31561 CCGGTTCCGA CCGGATCAG GCGGCCATCG CCGCGCGGA CCGCAGCGCC TGCATCAGGC
45 31621 GGGCCCGTGC GGACACCAGC CTGCACGCGT CCTCCAGGGA CCAGACGCGC GCGACGTACG
31681 CCGCGGCCAG CTCGCCGATC GAATGGCCCA CGAAGGCCTC CCGGCGTACG CCGCAGCGCT
31741 CGAGCTGTGC GCCGAGTGCG ACCTGGAGCG CGAACACCGC GGGCTGGGCG TACCGGGTGT
31801 CGTGAGGTC GAGCCCGCGG GGCACGTGCA GGGCGTCCAG CACCTCGCGG CGAGTGC3GG
31861 CGAAGACGTC GTAGGCGGCG GCCAGTCCGT CGCCCAT3CC GGGACGTTGT GAGCCCTGTC
50 31921 CGGAGAAGAG CCACACGAGG CCGCGGTCCG GTTCTCCGCG GCGGGTGACC GTGTCCGTGC
31981 CGATCAGCGC GGCCCGGTGC GGAAGGCCCG TCGGGGCGAG CAGGGCCGCG GCCACGCGC
32041 GCTCGTCTCT CTCGCCGGTG GCGAGGTGGG CGCGCAGCGG GTGTACCTGT GCGTCA3TG
32101 CCTGCGGGGT GCGTCCCGAG AGCAGCAGGG GCAGCGCTCC GGTGTGCGGT SCCGGGGCGG
32161 GTTCCGGGGC CCGTCCGGGG TGGCTTTCGA GGATGATGTG AGCTTTGGTG CCGCTAACCG
55 32221 CGAAGGAGGA CACCCCGCGG CCGCGTGGGC GGTGCTTTTC GGGCCAGGGG CCGGCGTGGG
32281 TGAGGAGTTC GACGGCGGCC GCGGTCCAGT CGACGTSCGA GGACGGCGTG TCCACGTCCA
32341 GGTGCGCGCG CAGGGTGCCG TGGCGATGG CGAGGACCAT CTTGATGACA CCGGCGACCG
32401 CCGCGCGCGC CTGAGTGTGG CCGATTTGG ACTTCA3GGA GCGGCGAGC ACCGGGGTGT
32461 CCGGATGCTG CCGGTAGGTG GCCAGTACCG CCTGCGCTC GATGGGGTGC CCGAGCTGG
60 32521 TCCCGGTGCC ATGCGCCTCG ACAGCGTCCA CATCCGCGCG GGTGAGCCCG GCGTTGGCCA
32581 GCGCCTGCCG GATCACCCGC TCCTGCGACG GCGCGTTCGG CCGCGACAAC CCGTTGGAAG

SUBSTITUTE SHEET (RULE 26)

32641 CACCGTCTTG GTTGACCGCC GAACCACGCA CGACCGCCAG GACATTGTGG CCGTGCCGCT
 32701 CGGCGTCCGA GAGCCTCTCG ACGATCAGCA CACCGGATCC CTCGGCGAAA CCGGTGCCAT
 32761 CAGCCGCATC CGCGAACGCC TTGCAGCGGC CGTCCGGGGA GAGGCCCGCG TGCTGGGAGA
 32821 AGTCCACGAA GCCGGACGGC GAGGCCATCA CCGTGACGCC GCGGACCAAG GCGAGCGAGC
 5 32881 ACTCCCGCGA GCGCAGCGAC TGCCCGGCTT GGTGCAGCGC CACCGAGCGAC GACGAACAGC
 32941 CCGTGTCCAC CGTGACCGCC GGACCCTCCA AACCCTAGAA GTACGACAGC CGACCGGACA
 33001 GCACACTGGT CTGGGTGCTG GTGGACCGA AACCGCCGCG CTCGGCTCCA GTGCCGTACC
 33061 CGTAGAAGTA GCCGCCCATG AACACGCCGG TGTCCTTCC SCGCAGCGAC TCCGGGAGGA
 33121 TCCCGGCGTG TTCCAGCGCC TCCCACGAGG TCTCCAGGAC CAGACGCTGC TGCGGGTCCA
 10 33181 TCGCCAGCGC CTCACGCGGA CTGATCCCGA AGAACGCCGC GTCGAAGTCC GCCACCCCGG
 33241 CGAGGAAGCC ACCATGACGC ACGGTCGACG TGCCCGGATG ATCCGGATCG GGATCGTACA
 33301 GCCCGTCCAC GTCCCAACCA CGGTCCGTCG GAAACGCCGT GATCCCGTCA CCACCCGACT
 33361 CCACAGCGCG CCACAAGTCC TCCGGCGAGC CGACCCACC CGGCAGCCGG CAGGCCATCC
 33421 CCACGATCGC CAACGGCTCG TCCGCGCGA CGGCGCGGGT CGCGGTACGC CGCGGGTGG
 15 33481 TGGCCCGCGC GCCGGCCAGT TCGTCCAGGT GGGCGGCGAG CGCCTGCGCC GTGGGGTGGT
 33541 CGAAGACGAG CGTAGCGGGC AGCGTCAGGC CCGTCGCGTC GGCCAGCCGG TTGCGCAGTT
 33601 CGACGCCCGT CAGCGAGTCG AAGCCCACTT CCCTGAACGC GCGCGCGGGT GCGATGGCGT
 33661 GGGCGTCCGC GTGGCCGAGC ACCGCGGCAG CGCTGGTACG GACGAGGTCG AGCATGTCGC
 33721 GCGCGGCCGG AGGTGCGGAG GTGCGCCGGA CGGCGGCAC GAGGGTGCCT AGGACCGGCG
 20 33781 GGACCGGTC GGACCGGCG ACGGCGGCGA GGTGAGCCG GATCGGCACG AGCGCGGGCC
 33841 GGTCGGTGTG CAGGGCCGCG TCGAACAGGG CGAGCCCTG TGCGCCGCTC ATCCGGGTCA
 33901 TGCCGTTGCG GCGGATGCGG GCCAGGTCGG TGSCSGTCAG CCGCCCGCCC ATCCCGTCCG
 33961 CCGCGTCCCA CAGTCCCCAG GCGAGCGAGA CGGCGGGCAG CCCCTGGTGG TGCCGGTGGC
 25 34021 GGGCGAGCGC GTCGAGGAAC GCGTTGCCGG TCGCGTAGTT GGCCTGACCC GCGCCGCCGA
 34081 ACGTGGCGGA TATGGACGAG TACAGGACGA ACGCCGCCAG GTCGAGATCG CGCGTCAGCT
 34141 CGTGCAAGTG CCAGGCGAGC TCCGCCTTGA CCCGCGACAC GCGTCCAC TGCTCCGGCC
 34201 GCATGGTCGT CACGGCCGCG TCGTCGACGA TCCCGGCCAT GTGCACGAC GCGCGCAGCC
 34261 GCTGGGCGAC GTCGGCGACG ACTGCGGCGA CTCGTCCGCG GTCGACGACG TCGGCGGCCA
 34321 CGTACCGCAC GCGGTCTGTC TCCGGCGTGT CGCCGGGCCG GCCGTTCGGG GACACCACGA
 30 34381 CGACCTCGGC GGCCTCTGTC ACGGTGAGCA GGTGGTCCAC GAGGAGGCGG CCGAGCCCGC
 34441 CCGTGCCGCC GGTGACGAGG ACGGTCCCGC CCGTCAAGCG GGAGGTTCCG GTGGCCCGCG
 34501 CGACACGGCG CAGACGGGCC GCACGCGCTG TGCCGTCCGC GACCCCGACG TCGGCTCGT
 34561 CGCCGCGCGC GAGCCCGGCC GCTATGGCGG CGGGCGTGAT CTCGTCCGCT TCGATCAGGG
 34621 CGACGCGGCC GGGATGCTCC GTCTCCGCGG TCCGACCAG GCCGCCGAGC GCTTCTCGC
 35 34681 CCGGATCGCC GGTACGGGTG GCCACGATGA GCCGGGATCG CGCCAGCGC GCTCGGCGA
 34741 GCCAGGTCTG CACGGTGGTG AGCAGGTCGC GGCCAGCTC CCGGTCCGG GCGCCGGGCG
 34801 AGGTGCCCGG GTCGCCGGGT TCCACGGCCA GGACCACGAC CGGGGGGTGC TCGCGTCCG
 34861 GCACGTCGCG GAGGTACGTC CAGTCGGGGA CCGGTGACGC GGGCACGGGC ACCCAGGCGA
 34921 TCTCGAACAG CGCCTCGGCA TCGGGGTGCG CGGCCCGCAC GGTCAGGCTG TCGACGTCAA
 40 34981 GGACCGGTGA GCCGTGCTCG TCCGTGGCGA CGATGCGGAC CATGTCCGGG CCGACGCGTT
 35041 CGGTCAGCAC GCGCAGCGCG GTGCGGCGC GCGCGTGGAT CCTACGCGG CCGTCCGCGA
 35101 ACCCCAGCCG GCGCCGCTCC GGGTCCGTGA AGACCGTCCC GAGGGCGTGC AGGGCCGCGT
 35161 CGAGCAGCAC GGGGTGCAGC CCGTACCGGG CGTCCGTGAG CTGTTCCGGC AGGCGGACCG
 35221 ACGCGTAGGC GCGGCCCTCC CCGTCCACA TCGCGGTGAT GGCCCGGAAC GCGGGCCCGT
 45 35281 ACGAGAGCGG CAGCGCGTCG TAGAAGCCGG TCAGGTCCGC CCGGTCCGGC TCGGCGGGCG
 35341 GCCAGTCCAC GGGCTCCGCC GGACCGCCAG TGTCACGCT CAGCGCTCCG GTCGCACTGA
 35401 GCGCCACGGG GCGCGTGCCG GTACGGCTGT GCAGACTCAC CGACCGCCGT CCGGACACCT
 35461 CCGTTCGAC GGTGGCTGG ATCTCCGTGT CGCGTCCGC GTCGACCACC ACCGCGCGA
 35521 CGATGCTCAG CTCCGCGATC TCCGGCGTGC CGAGCCGGGC TCCCGCTTCG GCGAGCCAGG
 50 35581 CCACGAGCGC CGAGCCGGGC ACGATGACCC GGCCGTCCAC CTCGTGGTCG GCGAGCCAGG
 35641 GCTGACGGCG TACCGAGACA CCGCGGTGCG CAGCGCGCCC TCGCCGTCCG GCGAGGTCCA
 35701 CCCACGAGCC GAGCAGCGGG TGGCCGGACG TTCCCGCGGG TTCCCGCTCG ATCCAGTAGC
 35761 GGTACCGCGG GAACGGGTAC GTGGGCAGCG GCACACCCG ACGCGTCGCG AACGACCAAG
 35821 TGACGGGCGC GCGCCGGACC CAGAGCGCG CGAGCGACCG AGTGAAGCGG TCCAGGCGC
 55 35881 CCGCGCGTCC CCGCAGTGTG CCGGTGACGA CCCTATGCGC ATGCGCGCGG ACGGTGCTCT
 35941 CCAGTGGCGT GGTGAGCACG GGATGCGCGC TGACCTCGAC GAACGCGCGG TATCCGCGGT
 36001 CCGCCAGGTG GCGGTCGCG GCGGCGAAC GAACGGTGG GCGCAGGTTG TCGTACAGT
 36061 AGGCGCGCTC CCGGGGCCCG TCCAGCCAG CCTCGTCCAC GGTGAGAGG AACGGGACCT
 36121 CCGGCGTCCG CCGAGTGATG CCGGCGAGAG CGTCGAGCAG CGCGCCGCGG ATCGTTTCCA
 60 36181 CATGCGCGGT GTGCGACCG TAGTCGACCG CGATCCGGCG GCGCGGGGGG GTGGCGGCCA

SUBSTITUTE SHEET (RULE 26)

36241 GCAGCTCCTC CACGGCGTCG GCCGCACCGG CGACAACGAT CGACGCGGGT CCGTTGACCG
 36301 CGGCGACCTC CAGGCGCCCG GCCCACACGG CGGCGTCGAA GTCGGCGGGC GGCACCGAGA
 36361 CCATGCCGCC CTGCCCGGCC AGTTCGGTGG CGACGAGTCG GCTGCGCACC GCGACGACCT
 36421 TCGCGGCGTC GTCCAGGGTG AGCACCCCGG CGACGAGGGG CGCGGGGACT TCGCCCTGGG
 5 36481 AGTGGCCGAC GACCGCGGCC GGGGCGACCC CGTGCGCACG CCACAGCTCC GCCAGCGCCA
 36541 CCATCACCGC GAACGACCGG GGCTGCACGA CATCGACCCG GTCGAACCGG GGCCTCCCGG
 36601 GCGGCTGGGC GATGACGTCC AGCAGGTCCC ATCCGGTGTG CCGGGCGAGC GCGGTGGCGC
 36661 ACTCGCGGAG CCGCCGGGCG AACACGGGGT CCGGTGGCGAG CAGTTCGGCA CCGATGCGCG
 36721 CCGACTGGGA GCCCTGCCCG GGGAACGGGA ACACGACACG TGTGTGGGTG ACCTCGCGCG
 10 36781 TCGCGGTCAC GGCCTCCGCG ACTTCGGCAC CACGGGCGAA CGCTCCGCG TCTCGGGCGG
 36841 CCACGACCGC CCGGTGGCGC ATGGCGCTCC GGGTGGTGGG GAGCGAGTGG CCGACCGCGG
 36901 CCGCGGCGCC AGTGAGCGGG GCCAGCTCTC CCGCGACGTC CCGCAGTCCC TCGGGGTCC
 36961 GCGCGACAT CCGCCAGACC ACGTCTCTCG GCACCGGCTC GGCTTCGGGT GCGGACACGG
 37021 GTGCGGGGCG GCGGGGGGGC CCGGCCTCCA GGACGACATG GCGTGTGGTG CCGCTGATGC
 15 37081 CGAACGACGA GACACCCGCA CGCCGGGCGC GCCCGGTGAC CCGCCACGGC TCACTGCGGT
 37141 GCAGCAGCCG CATGTGCGCG TCCAGTCCA CGTGCCGGGA CCGCTCGTCG ACCTGCAGCG
 37201 CCGCGCGCAG GACGCGGTGC CGCATCGCCA TGACCATCTT GATGACGCGG GCGACGCGCG
 37261 CCGCGGCGCT GGTGTGGCCG ATGTTGCACT TGAGCGAGCC GATCAGCAGC GATCAGCAGC
 37321 GTTCGCGCCC GTAGGCCACT TGCAGGGCCT GGGCCTCGAC GGGGTGCGCG AGACGGGTGC
 20 37381 CCGTGCCGTG TGCTCCACG GCGTCGACGT CACCCGGCGC CAGGCCGGCG TCGGCGAGCG
 37441 CACGCTGGAT GACGCGCTGC TCGCAGGCC CGTTCGGGGC GGACAGCCCG TTCGACGCGC
 37501 CCGCGGAGTT GACCGCGGAG CCGCGCACCA GCGCCAGCAC GGGGTGGCCG TGGCGGGTGG
 37561 CCGCGGAGAG CCGCTCCAGC ACCAGGACAC CGGCGCCCTC GGCGAAGCTC GTGCCGTCCG
 37621 CCGTGCTCCG GAAGGCCTTG GCACGGCCGT CGGGGGCGAG CCGCGCTGCG CCGGAGAACT
 25 37681 CGACGAACCC GGTGCTCGTC GCCATCACCG TGACACCGCC GACCGAGCCG AGCGAGCACT
 37741 CCGCGGAGCG CAGCGACCGC GCGGCCTGGT GCAGCGCCAC CAGCGACGAC GAACACGCGC
 37801 TGTGACGGT GACCGACGGG CCTCCAGAC CGAAGTAGTA CGAGAGCCCG CCGGAGAGAA
 37861 CCGTGGTCCG CGTGCCGGTC GCGCCGAAAC CGCCAGGTC CACGCCCCCG CCGTAGCCCT
 37921 GCGTGAACGC GCCCATGAAT ACGCCGGTGT CGCTGCCGCG GACGCTTTCG GGCAGGATGC
 30 37981 CCGCTCGTTC GAACGCCTCC CACGACGCTT CGAGGACCAG ACGCTGCTGC GGGTCCATCG
 38041 CCGAGCGCTC ACAGCGGCTG ATCCCGAAGA ACGCGGCGTC GAAGTCGGCG CCGCGGTGA
 38101 GGAAGCGGCC GTGACGCACG GAAACCTTGC CGACCGCGTC GGGGTTCGGG TCGTAGAGCG
 38161 CCGCGAGGTC CCAGCCGCGG TCGGCGGGGA ACTCGGTGAT CGCGTCCCCG CCGGAGTCGA
 38221 CCGAGCGCCA CAGGTCCTCC GGTGACCGCA CGCCACCGGG CATCCGGCAC GCCATGGCCA
 35 38281 CCGTCGCCAG CCGCTCGTTC CCGCCACCGG TCGGTGCGGG CACTGTGCGC GCGGAGCGCG
 38341 CCGGGGCGCG CTCACCCGCG CGTTCCTCAT CCAGGCGGGC GCGGAGCGCG GCGGTGTGCG
 38401 CCGGTGCGAA GACGGCGGTC GCGGAGAGCC GTACCCCGCT CGTCTCGGCG AGGTGTTGCG
 38461 GCAACCGGAC ACCGCTGAGC CCGGTCGATG CGAGGTCTCT GAACGCGCTG GTGGCGGTGA
 38521 TCTCGGAGGC GTCGGCGTGG CCGAGCACCG CCGCGGTGGC CGCACACACG ATGGCCAGCA
 40 38581 GGTACGATC GCGGTGCGCG TCGCGGTGCG GGTGTCTCTC CGCACGGGCG GCGATGCGGC
 38641 GGTGCGTCCG CTGCCGAGC GGCTCGGTGG GAATCGCCCG GACCATGAAC GGCACGTCCG
 38701 CCGCGAGGCT CCGTTCGATG AAGTGGGTGC CCTCGGCCTC GGTGAGCGGC CGGAACCCGT
 38761 CCGGACCCCG CTGCCGTCG GCGTCTGCAA GTTGTCCGGT GAGGGTGTCT GTGGTGTGCC
 38821 ACATGCCCCA GCGGATGGAG GTGGCGGGTT GGCGAGGGT GTGGCGGTGG GTGGCGAGGG
 45 38881 CCGTCAGGAA GCGGTGGCG GCGGCGTAGT TTCCTTGTCC GGGGTGCGCG AGGACGGCGG
 38941 CCGCGCTGGA GTAGAGGACG AAGTGGGTGA GGGGTGGTT TTGGGTGAGG TGGTGCAGGT
 39001 CCGAGGCGGC GTTGGCTTTG GGGTGGAGGA CCGTGGTGAG GCGGTGCGGG GTGAGGGCGT
 39061 CCGAGGATGC GTCTGCGAGG GTGGCGGCGG TGTGGAAGAC GCGGTGAGG GGTGGGGGA
 50 39121 TGTGGGCGAG GGTGGTGGCG AGTGGTGGG GGTGCGCGAC GTGCGAGGGG AGGTGGGTGC
 39181 CCGGGGTGGT GTCGGGGGT GGGGTGCGCG AGAGGAGGTA GGTGTGGGGT TGGTTCAGGT
 39241 GCGGGGCGAG GATGCCGCG AGGGTGGCG AGCCGCGGT GATGATGATG GCGTGTTCGG
 39301 GGTGAGGGG GGTGGTGGTG GGTGGGTGG TGGTCTGAG GGGGTGAGG TGGGTGCGT
 39361 GAGGGGTGTC GTGGGTGAGG CCGAGGTGG GGTGCTGAG GGTGGCGAGT TGGGCCAGGG
 39421 GAGGGGAGT GTGGGGGTGG TCGGTTTCCA TGAGGCGGAT GCGGTGGGGG TGTTCGTTCT
 55 39481 GCGCGGTGCG GTGAGGCGG GTGACGCTGG CCGCGGCGGG GTCGGTGGTG GTGTGACGA
 39541 GAGGGGTGTC GTCGGTGGTG GTGAGGTGCT GTTGACGGG GTTCAGGACG CCGGTGGCGC
 39601 GGTGTGGGG CCGGTGGGT ATGCTCTCG GGTGCTCGCG GTGGGCGGCG GTGATCAGGA
 39661 GGTGCTCTC CCGGAGGTCA CCGCTCTAGA CCGCTCTGGC GACCGCGAGC CACTCGAACC
 39721 GAGCGGGTT CCGCCCCGAC GGGGTGTGG CCGCTCTCT CAGCACCAGC GAGTCCACCG
 60 39781 ACACGACAGG ACGGCCATCC GGTGCGGCA CCGCGACGGC GACGCGGGC TCCCCCGGG

SUBSTITUTE SHEET (RULE 26)

39841 TTAGGGGCGAC GCGCACCGCG GCGGCCCGCG TGGCGTTGAG GCGCACGCCC GTCCAGGAGA
 39901 ACGGCAGCTC GATCCCGCCG CCCGCGTCGA GGCGCCCGGC GTGCAGGGCC GCGTCGAGCA
 39961 GTGCCGGATG CACACCGAAA CCGTCCGCCT CGGCGGCCTG CTCGTCGGGC AGCGCCACCT
 40021 CGGCATACAC GGTGTCAACA TCACGCGAGG CAGCCCGGAA CCCCTGGAAC GCGGACCCCT
 5 40081 ACTCATAACC GGCATCCCGC AGTTCGTCTAT AGAACCCCGA GACGTCGACG GCGCGCGCCG
 40141 TGGCCCGCGG CCACTGCGAG AACGGCTCAC CGGAAGCGTT GGAGGTATCC GGGGTGTGCG
 40201 GGGTCAGGGT GCCGCTGGCG TGCCGGGTCC AGCTGCCCGT GCCCTCGGTA CGCGCGTGGA
 40261 CGGTCACCGG CCGCCGTCCG GCCTCATCGG CCCCTTCCAC GGTACCCGAC ACATCCACCG
 10 40321 CTGCGGTACG CGGCACCACG AGCGGGGATT CGATGACCAG TTCATCCACC ACCCGCAAC
 40381 CGGTCTCGTC ACCGGCCCGG ATGACCAGCT CCACAAACGC CGTACCCGGC AGCAGAACCG
 40441 TGCCCCGCAC CGCGTGATCA GCCAGCCAGG GATGCGTACG CAATGAGATC CGGCGCGTGA
 40501 GACAAACACC ACCACCGTCG TCGGCGGGCA GTGCTGTGAC GCGCGCCAGC ATCGGATGCG
 40561 CCGCCCGCGT CAGCCCGGCG GCGGACAGGT CGGTGGCACC GGCCGCCTCC AGCCAGTACC
 15 40621 GCGTGTGCTC GAACGCGTAG CAGCAGCCG CCGCGGACCG CCCCGGACCG GGTTCGACCA
 40681 CGGTGCCCCA GTCCACCCCG GCACCCAGAG TCCACGCCTG CGCCAACGCC CCCAGCCACC
 40741 GCTCCAGGCC ACCGTCAACA GTCCGCAACG ACGCCACCGT GCGGGCCTGT TCCATCGCCG
 40801 GCAGCAGCAC CGGATGGGCA CTGCACTCCA CGAACACCGA CCCGTCCAGC TCCGCCACCG
 40861 CGCATCCAG CGCGACAGGG CGACGCGAGT TCCGGTACCA GTACCCCTCA TCCACCGCT
 20 40921 CGGTACCCCA GCGGCTGTCC ACGGTGCAAC ACCACGCCAC CGACCCGGTC CCGCCGGAAG
 40981 TCCCTTCAG TACCTCAGCG AGTTCGTCTT CGATGGCCTC CACGTGAGGC GTGTGGGAGG
 41041 CSTAGTCGAC CGCGATACGA CGCACCCGCA CCCATCAGC CTCATACCGC GCCACCACCT
 41101 CCTCCACCGC CGACGGGTCC CCCGCCACCA CCGTCGAAGC CGGACCATTA CGGCGCGCA
 41161 TCCACACACC CTCGACCAGA CCCACCTCAC CGGCCGGCAA CGCCACCGAA GCCATCGCCC
 25 41221 CCGCGCCGGC CAGCCGCGCC GCGATCACCC GACTGCGCAA CGCCACCACG CGGGCGGCGT
 41281 CCTCCAGGCT GAGGGCTCCG GCCACACACG CCGCCGCGAT CTCCCCTGCG GAGTGTCCGA
 41341 CCACAGCGTC CGGCACGACC CCATGCGCCT GCCACAGCGC GGCCAGGCTC ACCGCGACCG
 41401 CCGAGTGGC CGGCTGGACC ACCTCCACCC GCTCCGCCAC ATCCGACCGC GACAACATCT
 41461 CCGGCACATC CCAGCCCGTG TCAGGCAACA ACGCCCGCGC ACATCTCTCC ATACGAGCCG
 30 41521 CGAACACCGC GGAACGGTCC ATGAGTTCCA CGCCCATGCC CACCCACTGG GCACCTGCGC
 41581 CGGGGAAGAC GAACACCGTA CGCGGCTGAT CCACCGCCAC ACCCATCACC CGGGCATCAC
 41641 CCAGCAGCAC CGCACGGTGA CCGAAGACAG CACGCTCAGC CACCAACCCC TGCGCGACCG
 41701 CGGCCACATC CACCCACCCC CCGCGCAGAT ACCCTTCCAG CCGCTCCACC TGCCCCCGCA
 41761 GACTCACCTC ACCACGAGCG GACACCGGCA ACGGCACCAA CCCATCACCA CCCGACTCCA
 35 41821 CACGCGACGG CCCAGGAACA CCTTCCAGGA TCACGTGCGC GTTCGTACCG CTCACCCCGA
 41881 ACGACGACAC ACCCGCATGC GGTGCCCCGAT CGGACTCGGG CCACGCGCTC GCGTGGTGA
 41941 GCAGCTCCAC CGCACCGGCC GACCACTCCA CATGCGACGA CGGCTCGTCC ACGTGCAGCG
 42001 TCTTCGGCGC GATCCCATGC CGCATCGCCA TGACCATCTT GATGACACCG GCGACACCCG
 42061 CAGCCCGCCTG CGCATGACCG ATGTTGCGACT TGACCGAACC GAGGTAGAGC GCGGTGTGCG
 40 42121 GGTCTGCCC GTAGGCGCGG AGGACGGCCT GCGCCTCGAT CCGGTGCGCC AGCGCGTGC
 42181 CGGTGCGGTG CGCCTCCACC ACGTCCACAT CGGCGGCGCG CAGTCCGGCG TTGACCAACG
 42241 CCTGCGGAT CACGCGCTGC TGGCGGACGC CGTTGGGGGC GGACAGTCCG TTGGAGGCAC
 42301 CGTCTGGTT CACCGCCGAG CCGCGGACGA CCGCGAGAAC GGTGTGCCCG TTGCGTCCG
 42361 CGTCGGAGAG CCGCTCCAGC ACGAGAACGC CGACGCCCTC GGCGAAGCCG GTCCCGTCCG
 42421 CCGCGTCGGC GAACGCCCTT CACCGTCCGT CCGGGGAGAG TCCGCGCTGC CCGGAGAACT
 45 42481 CCACGAGCTC TGCGGTGTTT GCCATGACGG TGACACCGCC GACCAGCGCC AGGGAGCACT
 42541 CCGCGGCCCC CAGTGCTGTG GCCGCTGGT GCAGGGCGAC CAGCGACGAC GAGCACGCCG
 42601 TGTGACCGCT GACCGCCGGG CCCTGAAGTC CGTACACGTA CGAGAGGCGC CCGGACAGGA
 42661 CGCTCGTCTG CGTCGCGGTG ACACCGAGCC CGCCAGGTG CCGGCGGACG CCGTAGCCCT
 50 42721 GGTGGAACGC GCGCATGAAC ACGCCGGTGT CGCTCTCCCG GAGCCTGTCC GGCACGATGC
 42781 CGGCGTTCTC GAACGCCTCC CAGGAGGTCT CCAGGATCAG GCGCTGCTGG GGGTCCATCG
 42841 CCAGCGCCTC GTTCGGACTG ATGCCGAAGA ACGCGGCGTC GAACCCGGCG CCGGCCAGGA
 42901 ATCCGCGCTG GCGTGTGCTG GAGCGGCGCG CCGCGTCCCG GTCCGGGTGCG TACAGCGCGT
 42961 CACGTCCTCA GCGCCGGTCC GTGGGGAAC CCGTGATCGC CTCGGTACCG GCGGCGACGA
 55 43021 GCGGCCACAG GTCTCTCGCG GAGGCGAACC CCGCGGGCAG TCGGCACGCC ATGCCGACGA
 43081 TCGCGACGGG GTGCGCGGAG CCGGCGGTCT GCGCGGTGCG GGTGCGCGCT GTCGCGGAGC
 43141 CCGCGAGGTG GCGGCGGAAC GCACGCGGAG TCGGGTGGTC GAACGCGGTT GACGCGGGCA
 43201 CCGCGAGACC CCGCGCGCGC GCGAGGCTGT TCGTGAACCT GACGGTGGTG AGCGAGTCCA
 43261 GCGCGTCTC GCGGAACGTC CCGTCCCGCG ACAGTGTGTC GCGCGCGCGC AGGCGGAGGA
 43321 CGGTGGCGAC GCTGTGCGCG ACCAGGTGCA GCGGTACGTC CTCCCGGCCC GCACGGGCGG
 60 43381 CCGCGAGGCG GTTCGCCCCA TCGTGTTCGG TGGCGTCGGG CTCGGCCGGT CCGGTACGTG

SUBSTITUTE SHEET (RULE 26)

43441 CGGTGAGGAT GGGCGGCGTG GCGCCCGCCA TCGTCCGCGC CCGCGCCCGG GCGGAACCGG
 43501 TCCGGGCGAC GATGTACGAG CCGCCGCGCG CGATGGCCTT CTCGATCAGG TCGCCGGTGA
 43561 GCGCCGGCGG TTCGATGCCG GGCAGCGCGC GGACGGTGAC GGTGGGGAGT TCGACGAGCG
 43621 CCGGTGGCGG GGTGTGGCGG TCGGCGCGCG CCGGGCGGTC GAGCAGGACG TCGACGAGCG
 5 43681 CGCCGGGGTT CGCGGCTTCC TCGGCTGCGG TGGTCACGTG GGTGAGGCGG GTCTCGTCGC
 43741 GGAGCAGGCC GGCACGCGTG TCGGCGTCTT CCCCCTGAC CAGGACCGGC GCGTCCGGGC
 43801 CGATCGGAGG CCGCACGCGT AGGACCATCT TGCCGGTGTG CCGGGCGTGG CTCATCCAGC
 43861 CGAACGCGTC CCGCGCACGG CGGATGTCCC ACGGCTGCAC CCGCAGCGGG CACAGCTCAC
 10 43921 CGCGGTGCGA CAGGTGAGG ATGGGCACGG AGGATCTCCG CAGGCGCGCG GATCCACGT
 43981 CGGCCAGGTC GAACGGCTGC TGGGCGGCGT GCGGATGTC GGTCTTGCCC ATCTCGACGA
 44041 ACCGGCCGCG CCGTGCGAGC AGGCCGATGG ACGGCTCGAG GAGTTCACCG GTGAGCGAGT
 44101 TGAGCACGAC GTCGACCGGC GGGAAGGTGT CCGCGAACGC GCGCTGCGG GAGTTCGCCA
 44161 CATGGTCGCT GTCGAAGCCG TCGGCGTGCA GCAGGTGTTG TTTGGCGGGA CTGGCGGTGG
 44221 CGTACACCTC GCGCCGAGG TGGCGGGCGA TCCGGGTGCG CGCCATGCCG ACACCGCCCG
 15 44281 TCGCGGCGTG GACAGGACC TTCTGGCCGG GTCGAGCTC GCGCGCTCG ACGAGGCGGT
 44341 ACCAGGCGGT GCGGAACACG ATGGGCACGG ACGCGCGAT CCGGAACGAC CATCCCCGTG
 44401 GGATCCGTGC GACAGCCGC CCGTCCGCGA CCACGCTGCG CCGGAACGAC TCGTGCAGT
 44461 GACCGAACAC GCGGTGCGCG GGGGCCAGGT CGTCGACGCC GGGTCCGACT TCGGTACGCA
 20 44521 TGCCCGCGGC CTCGCCGCC ATCTCGCCCT CCGCCGGTA GGTGCCGAGC GCGATCAGCA
 44581 CGTCGCGGAA GTTCAGCCCC GCGGCGCGGA CGTCGATGCG GACCTCGCCG GCGGCCAGGG
 44641 GCGCGGCGGG ACGTCGAGCG GCGGACGAC GAGGTCGCGG AGCGTCCCG AGGCGGCGCG
 44701 GCGCAGCGCG CACTGCGCGG TCGGCGAGG GGTGGTGTG CCGCGGTACC AGCCGGGGCA
 44761 CGTAGGCCAC GCGGCGCGC AGCGGATCT GGGGTTCGCC GAGCGAGGC GCGGCGGGGA
 25 44821 CGAGGTGCTC ATCGCCGTCC GTGTCCACCA GCACGAACGA TCCGGGTTCG CCGGCTGCG
 44881 GCGCGAGCGC CTCGTCCAG AGCCGGGCTT GGTCCGCGTC CCGGATCTCG GCGGCGCGCA
 44941 CGCCACCGCG GCGGCGGGTG ACGACCGTCC GCGGGGTGA CCGGGTGGCG GCGAGGTGCG
 45001 GCGGCTCCCA GACAGTTTCG CACAGCGTGG CCTCGCCACT GCGGGTGGCG ACCAGATGGG
 45061 CCGGACGCCC CCGAGCCCG GCGCGCTGGA CTTTGCCCGA CCGGTGCGG GGGATCGTGG
 30 45121 TGACGTGCCA GATCTCGTCG GCGACCTTGA AGTAGGCGAG CCGGCGCGCG CACTCGGCGA
 45181 GGATCGCCTC GCGGGGACG GCGGGGCGGT CCGAAACGAC GTAGAGCAGC GGTATGTCG
 45241 CGAGGACGGG GTGCGGGCGG CCGCGCGCGG CCGGCTCCCG GACACCGGCC ACCTCCTGGG
 45301 CGACGGTCTC GATCTCCCG GGGTGGATGT TCTCCCCGCC GCGGATGATC AGCTCCTTGA
 45361 CCGGCGCGGT GATCGTCACG TGTCCGGTCT CCGCCTGACG TGCGAGGTCC CCGGTGCGGT
 35 45421 ACCAGCCGTC CACGAGCACC TGGGCGGTG CCTCCGGCTG GCGGTGGTAG CCGAGCATGA
 45481 GGTGCGGCC GCTCGCCAC AGCTCGCCCT CCGCGCGG TCGCACGTG GCGCGGACA
 45541 CCGGGTCGAC GAACCGCAGC GACAGGCGCG CCGCGGCG CCGCACGAC CCGGGAACCC
 45601 GCGCATCCTC CAGGGTGTG GCGGTGAGCG AGCCGGTCTT CTCGGTGCAG CCGTACGTGT
 45661 CGAGCAGGGG CACGCCGAAC GTCGCTCGA AATCCCTGGT GAGCGACGCC GCGAGGTGG
 45721 ATCCGGCGAC CAGCGCCAGC CGCAGCGCGC GAGCCCGCG CCGCGCGGAC ACGGCGCGCA
 40 45781 GGAGGTAGCG GTACATCGTC GGCACGCGCA CGAGCACGGT GCTGGAGTGT TCGGCCAGGG
 45841 CGTCGAGGAC GTCACGCGCG ACGAAGCCGC CCAGGATACG GCGGACGCG CCGACCGTGA
 45901 GGAGCGGAG CAGGCGAGG TGGTGGCGCA GGCTGTGGAA CAGCGGGGCG GCGCAGAGCA
 45961 GTTCGTGCTC CTCGGTCAGC CGCCAGGACG GCACGTGCGA GTGCATCGCG GACCAAGGC
 46021 CGCTGCGCTG TCGGGAACCC ACGCCCTTGG GACGGCCGGT GGTGCCGAG GTGTAGAGCA
 45 46081 TCCAGGCGGG TCGTCCAGG CCGAGGTCTG CCGGGGCGG GCACGGCGGC TCGGTCCCGG
 46141 CGAGGTCTCT GTAGGAGACG CAGTCCGGTG CCGGCGCGC GACGAGCACG ACGGTGGCGT
 46201 CCGTGCCGGT GCGGCGCACC TGGTCGAGGT GGGTTTCGTC GGTGACCAGC ACGGTGCGCG
 46261 CCGAGTCCGT CAGGAAGTGG GCGAGTTCG CCGGCGCGG GTCCGGGTG AGCGGGACCG
 50 46321 CGACGGCGG GCGCGGGCG GCGGCGAGGT AGACCTCGAT GGTCTCGAT CCGTTGCCGA
 46381 GCAGCATCGC GACCCGCTCG CCGCGGTGCA CCGCGGACG GCGAGGTGT CCGGCGAGCC
 46441 GCGCGGCGCG GAGCCGGAGT TCGGTGTACG TCACGGCGCG TTGGGAATCC GTGTAGGCGA
 46501 TCGGTCGCG CCGTCCCTCG GCATGGATGC GGAGCAATTC GTGCAACGGC CCGATTGGTT
 46561 CCACACGCGC CATGGAAACA CTTTCTCTC GACCAACCGC ACAACAGCAC GGAACCGGCG
 55 46621 ACGAGTAGAC GCGGCGCAGC CTAGCAGCGT TTTCCGGACC GCGACCCCTT GAAGATCCCT
 46681 CTACCGTGCC CCGGCTCCCG GGACGCTCAT CTAGGGGGTT GCGCGCATAC CCGGTGCGT
 46741 AATTGCTTCT CTGATGACCG ATGCGGAGC CAGGGGAAGG GTGGAGGCTT TGTCCATATC
 46801 TGTCACGCGC CCGTATTGCC GCTTCGAGAA GACCGGATCA CCGGACCTCG AGGGTGAAGA
 46861 GACGGTGTCT GCGCTGATCG AGCACGGCAC CCGCCACACC GAGGTGTGCG TGGTGGAGCG
 46921 TGCTCCCGCG ACCGCGGTGC ACACACGAC CCGTGACGAC GAGGCGTTCA CCGAGGTCTG
 60 46981 GCACGCACAG CCGCCTGTG AGTCCGGCAT GGACAACGGC ATCGCCTGGG CCGCACCGCA

SUBSTITUTE SHEET (RULE 26)

47041 CGCGTACCTG TTCGGTGTGCG TGCGCAACCGG CGAGAGCGGC AGGTACGCGG ATGCCACCGC
 47101 GGCCCTCTAC ACGAACGTCT TCCAGCTCAC CCGGTGCGTG GGGTATCCCC TGCTCGCCCG
 47161 GACCTGGAAC TACGTCAGCG GTATCAACAC GACGAACGCG GACGGGCTGG AGGTGTACCG
 47221 GGACTTCTGC GTGGGCCGCG CCCAGGCGCT CGACGAGGGC GGGATCGACC CGGCCACCAT
 5 47281 GCGCGCGGCG ACCGGTATCG GCGCCACCGG GGGCGGCATC ACCTGCGTGT TCCTCGCCCG
 47341 CCGGGCGGGA GTGCGGATCA ACATCGAGAA CCCC GCCGTC CTCACGGCCC ACCACTACCC
 47401 GACGACGTAC GGTCCGCGGC CCCC GTCTT CGCACGGGCC ACCTGGCTGG GCCCGCCGGA
 47461 GGGGGGCGCG CTGTTTATCT CCGCGACGGC CGGCATCCTC GGACACCGAA CGGTGCACCA
 47521 CCGTGATGTG ACCGGCCAGT GCGAGGTGCG CCTCGACAAC ATGGCCCGGG TCATCGGCGC
 10 47581 GGAGAACCTG CCGCGCCACG GCGTCCACG GGGGCACGTC CTCGCCGACG TGGACCACCT
 47641 CAAGGTCTAC GTCCGCCGCC GCGAGGATCT CGATACGGTC CGCCGGGTCT GCGCCGACG
 47701 CCTGTGAGC ACCGCGGCCG TCGCCCTTTT GCACACCGAC ATAGCCCGCG AGGATCTGCT
 47761 CGTCGAAATC GAAGGCATGG TGGCGTGACA ATACCCGGTA AAAGGCCCGC GACGCTGCGC
 15 47821 CTCGGCGGAT CCGCGAAGAG AAAGAAGAGC GTCACCGCAC AGCGCGGCAG CCCGGTCTTT
 47881 TCGTCTTCG CACAGCGGCG GATCTGGTTT CTCCAGCAAT TGGACCCGGA GAGCAACGCC
 47941 TATAATCTCC CGCTCGTGCA ACGCCTGCGC GGTCTATTGG ACGCGCCGGC CTTGGAGCGT
 48001 GCGCTGCGCG TCGTCTGTCG GCGCCACGAG GCGTTGCGGA CCGTGTTCGA CACCGCCGAC
 48061 GCGGAGCCCG TCCAGCGGGT GCTTCCCGCC CCGGAACACC TCCTGCGCCA CGCGCGGGCG
 48121 GGCAGCGAGG AGGACGCCCG CCGGCTCGTC CGCGACGAGA TCGCCGCGCC GTTCGACCTC
 20 48181 GCCACCGGCG CGTTGATCAG GGCCCTGCTG ATCCGCTCG GTGACGACGA CCACGTTCTC
 48241 CGCGTGACCG TGCACCATGT CGCCGCGCAC GGCTGGTCTG TCGGGCTCCT CCAACATGAA
 48301 CTCGCAGCCC ACTACACGGC GCTGCGCGAC ACTGCCCGCC CTGCCGAAC SCCGCCGTTG
 48361 CCGGTGCACT ACGCCGACTT CGCCGCTGCG GAGCGGCGCG AACTCACCGG CGCCGGACTG
 48421 GACAGGCGTC TGGCCTACTG GCGCGAGCAA CTCCGGGGCG CCGCGGCGCG GCTCGCCCTC
 25 48481 CCCACCGACC GTCCCCGCCG GCCGGTTCGCC GACGCGGACG CCGGCATGGC CGAGTGGCGG
 48541 CCGCCGCGCG CGCTGGCCAC CGCGGTCTCT ACGCTCGCGC GCGACTCCGG TGCGTCCGTG
 48601 TTCATGACCC TGCTGGCGCG CTTCCAAGCG GTCTTCGCCC GGCAGGCGGG CACGCGGGAC
 48661 TGCTGGTTCG GCACGCCCGT GGCGAACCGT ACGCGGGCGG CGTACGAGGG CCTGATCGGC
 48721 ATGTTCTGCA ACACGCTCGC GCTGCGCGGC GACCTCTCGG GCGATCCGTC GTTCGGGAA
 30 48781 CTCCTCGACC GCTGCCGGGC CACGACCACG GACGCGTTCT CCCACGCCGA CCTGCCGTTT
 48841 GAGAACGTCA TCGAACTCGT CGCACCGGAA CGCGACCTGT CCGTCAACCC GGTCTCCAG
 48901 GTGCTGTTGC AGGTGCTGCG GCGCGACGCG GCGACGGCCG CGCTGCCCGG CATCGCGGCC
 48961 GAACCGTTCC GCACCGGACG CTGGTTTACC CGCTTCGACC TCGAATTCCA TGTGTACGAG
 35 49021 GAGCCGGTG GCGCGCTGAC CGGCGAATG CTCTACAGCC GTGCGCTGTT CGACGAGCCA
 49081 CGGATCACGG GGTGCTGGA GGAGTTACG ACGGTGCTTC AGGCGGTAC CGCCGACCCG
 49141 GACGTACGGC TGTGCGGGCT GCCGGCCGGC GACGCGACGG CCGCAGCGCC CGTGGTGCCC
 49201 TCGAACGACA CCGCGCGGGA CCTGCCCGTC GACACGCTGC CCGGCCGTGCT GGCCCGGTAC
 49261 GCGCGACGCA CCCCCGCGCG CGTGGCCGTC ACCGACCCCG ACATCTCCCT CACCTACGCG
 49321 CAGCTGGAAC GCGGGGCGAA CCGCCTCGCG CACCTGCTCC GCGCGCGCGG CACCGCCACC
 40 49381 GCGCACTGCG TCGGGATCTG CGCCGATGCG GCGCGCGACC TGATCGTCCG CATCGTGGGG
 49441 ATCCTCAAAG CCGCGCGCGC TTATGTGCGG CTGGACCCCG AACATCTCTT GGAGCGCACG
 49501 GCGTTCTGTC TGGCCGACGC GCGGTGACG ACGGTGGTGG CGCACGAGGT TACCGTTCC
 49561 CCGTTCCCCG ATGTGCCGCA CGTGGTGGCG TTGGACGACC CGGAGCTGGA CCGGCAGCCG
 49621 GACGACACGG CGCCGACGCT CGAGCTGGAC CCGGACAGCC TCGCCTACGC GATCTACAG
 45 49681 TCCGGGTCGA CCGGCAGGCC GAAGGCCGTG CTCATGCCGG GTGTACGCG CGTCAACCTG
 49741 CTGCTCTGGC AGGAGCGCAC GATGGGCCCG GAGCCGGCCA GCCGCACCGT CCAGTTCTGT
 49801 ACGCCACGCT TCGACTACTT GGTGCAGGAG ATCTTTTCCG CGTGTCTGG CGGCACGCTC
 49861 GTCATCCCGC CGGACGAGT GCGGTTTCAG CCGCCGGGAC TCGCCCGGTG GATGGACGAA
 49921 CAGGCGATTA CCGGATCTA CGCGCCGACG GCCGTACTGC GCGCGTGTAT CGAGACGTC
 50 49981 GATCCGCACA GCGACAGCT CGCCGCCCTG CCGCACCTGT GCCAGGGCGG CGAGGCGCTG
 50041 ATCCTCGACG CGCGGTTGCG CGAGCTGTGC CCGCACCGGC CCCACCTGCG CGTGACAAT
 50101 CACTACGGTC CGGCCGAAAG CCAGCTCATC ACCGGGTACA CGCTGCCCGC CGACCCCSAC
 50161 GCGTGGCCCG CCACCGCACC GATCGGCCCG CCGATCGACA ACACCCGCAT CCATCTGCTC
 50221 GACGAGGCGA TCGGGCCGTT TCCGACGGT ATGCCGGGGC AGCTCTGCGT CGCCGGCGTC
 55 50281 GGCCTCGCCG GTGGGTACCT GGCCGCTCCC GAGCTGACCG CCGAGCGCTG GGTGCCGGGA
 50341 GATGCGGTGCG GCGAGGACG CATGTACCTC ACCGGCGACC TGGCCCGCGG CGCGCCGAC
 50401 GGCGACCTG AATTCTCTCG CCGGATCGAC GACCAAGTCA AGATCCGCGG CATCCGCGTC
 50461 GACCGCGCTG AGATCGAGAG CCTGCTCGCC GAGGACGCCG GCGTACGCA GCGGCGGCTG
 50521 TCCGTGCGCG AGGACCGGCG GGGCGAGAAG TTCCTGGCCG CGTACGTCGT ACCGGTGGCC
 60 50581 GGCCGCGACG GCGACGACTT CGCCGCGTGC CTGCGCGCGG GACTGGCCCG CCGGCTGCCC

SUBSTITUTE SHEET (RULE 26)

50641 GCGCGGCTCG TGCCCTCCGC CGTGGTCTTG GTGGAGCGAC TGCCGAGGAC CACGAGCGGC
50701 AAGGTGGACC GCGCGCGGCT GCGCGAGGCG GAGCGGGGCC CGGCGTCGAC CGGGGCGGTT
50761 ACGCCCCGCA CCGATGCCGA GCGGAGGCTG TGCCGGATCT TCCAGGAGGT GCTCGACGTC
50821 GCGCGGCTCG GTGCGGACGA CGACTTCTTG AGGCTCGGCG GCGACTCCCT GCTCGGCAAC
50881 GCGGTGCTCT CCGGCATCCG CGCCGAGGCT GGTGCGGATG TCCGCTGCG TACGCTCTTC
50941 GACGGGCGGA CGCCCGCCGC GCTCGGCTCT GCGGCGGACG AGGCCGGGCC GGCCGCCCTG
51001 GCGCGGATCG CGCCCTCCGC GGAGAGCGGG CCGGCCCCCC TCACCGCGGC ACAGGAACAG
51061 ATGCTGCACT CGCACGGCTC GCTGCTCGGC GCGCCCTCCT ACACGGTCGC CCGGTACGGG
51121 TTCCGGCTGC GCGGGCCACT CGACCGCGAA GCGCTCGACG CGGCACTGAC CCGGATCGCC
51181 GCGCGCCACG AGCCGCTGCG GACCGGGTTC CCGGATCGGG AACAGGTCGT CCGGCGCGCC
51241 GCTCCGGTGC GCGCGGAGGT GGTTCGCTTG CCGGTGCGCG ACGTCGACGC CGCGGTCCGG
51301 GTCGCCCCACC GGGAGCTGAC CCGGCCCTTC GACCTCGTGA ACGGGTCGTT GCTGCGTGCC
51361 GTGCTGCTGC CGCTGGGCGC CGAGGATCAC GTGCTGCTGC TGATGCTGCA CCACCTCGCC
51421 GTGACGGAT GGTCTTCGA CCTCCTGCTC CGGGAGTTGT CGGGGACGCA ACCGGACCTT
51481 CCGGTGTCTC ACACGGAGT GGCCCTGCTG GAACGGAGTC CGGCCGTGAT CGCGGCCAGG
51541 GAGAACGACC GGGCCTACTG GCGCCGCGCG CTGGGGGGCG CCACCGCGCC CGAGTGCACC
51601 GCGGTCCGGC CCGGCGGGGC ACCGACCGGG CCGGCGTTCC TGTGGACGCT CAAGGACACC
51661 GCGGTCTCTG CCGCACGCGG GGTGCGGAC GCGCACGACG CGACGTGCA CGAAACCGTG
51721 CTGCGGCGCT TCGCCCTGGT CGTGGCGGAG ACCGCGGACA CCGACGACGT GCTCGTCGCG
51781 ACGCCGTTTC CCGACCGGGG GTACGCGGCG ACCGACCACC TCATCGGCTT CTTGCGGAAG
51841 GTCCTCGCGC TGCGCTTCGA CCTCGGCGCG ACGCGTCTGT TCCCGAGGT GCTGCGCCGG
51901 GTGCACACCG CGATGGTGGG CGCGCACTCG CACCAAGCGG TGCCCTACTC CGCGCTGCGC
51961 GCGGAGGACC CCGCGCTGCC GCGGCGCGCG GTGCTGTTCC AGCTCATCAG CGCGCTCAGC
52021 GCGGAAGTGC GGTGCGCCGG CATGCACTAC GAGCGGTTCC CCGTCTGCGC CGAGACCGTC
52081 GACGAGATGA CCGGCGAAGT GTGATGACAC CTCTTCGACG ACGGTGCGAC CGTCTCCGGC
52141 GCGGTGGTCC ACGATGCCGC GCTGCTCGAC CGTGCCACCG TCGACGATTT GCTCACCCTG
52201 GTGGAGGCGA CGCTGCGTGC CGCGCGGCGC GACCTCACCG TACGCGTCAC CCGTTACGTG
52261 GAAAGCGAGT AGCCATGCCG GAGCAGGACA AGACAGTCGA GTACCTTCGC TGGGCGACCG
52321 CGGAAGTCCA GAAGACCCGT GCGGAAGTTC CCGCGCACAG CGAGCCGTTG CGGATCGTGG
52381 GGATGGCCTG CCGGCTGCCC GCGGGGGTTC CGTCCGCGGA GGACCTGTGG CAGTTGCTGG
52441 AGTCCGGTGG CGACGGCATC ACCGCGTTCC CCACGGACCG GGGCTGGGAG ACCACCGCCG
52501 ACGGTGCGCG CCGCTTCCTC ACCGGGGCGG CCGGCTTCGA CGCGCGGTTT TTCGGCATCA
52561 GCGCGCGCGA GCGGCTGGCG ATGGACCGCG AGCAGCGCCT GGCCCTGGAG ACCTCGTGGG
52621 AGGCGTTTCA GCACGCGGGC ATCGATCGCG AGACGCTGCG GGGCAGTGAC ACGGGGGTGT
52681 TCCTCGGCGC GTTCTTCCAG GGGTACGCGA TCGGCGCGCA CTTGACGGT TACGGGACCA
52741 CGAGCATTCA CACGAGCGTG CTCTCGGCGC GCCTCGCGTA CTTCTACGGT CTGGAGGGTC
52801 CCGCGGTGAC GTGCGACAG GCGTGTCTCT CGTCTGCTGG GCGGCTGAC CAGGCGGGGC
52861 AGTCGCTGCG CTCCGGCGAA TGCTCGCTCG CCCTGGTTCG CCGGCTGACG GTGATGGCCT
52921 GCGCGGCGGG GTTCGCGGAC TTCTCGGAGC AGGGCGGCGT GGCCCCCGAC GCGGCTGCA
52981 AGGCCTTCGC GGAAGCGGCT GACGGCACTG GTTTCGCGGA GGGGTCCGGC GTCTGATCG
53041 TCGAGAAGCT CTCCGACGCC GAGCGCACTG GCCACCGCGT GCTGGCGGTG GTCCGGGGTT
53101 CCGGCGTCAA CCAGGACGGT GCTTCCGACG GGCTGTCCGC GCGGAACGGG CCGTGCAGG
53161 AGCGGGTGAT CCGGCGAGGC CTGGCCGACG CCGGACTCAC CCGGCGGAC GTGGAGCCCG
53221 TCGAGGCCCA CCGCACCGGC ACCAGGCTGG GCGACCCCAT CGAGGACAG GCCGTGCTGG
53281 CCACCTACGG GCAGGGGCGC GACACCGCTG TGCTGCTGGG CTCGCTGAAG TCCAACATCG
53341 GCCACACCCA GCGCGCCGCG GCGCTCGGCG GTGTCATCAA GATGGTCTTC GCCATGCGGC
53401 ACGGCACCTT GCGCCGACG CTGCACTGG ACACGCGCTC CTCGACGTC GACTGGACGG
53461 CCGGCGCGGT CGAACTCCTC ACCGACGCGC GGCCCTGGCC CGAAACCGAC CGCCACGGC
53521 GCGCGGTTGT CTCTCTCTTC GCGGTGAGCG GCACCAACGC CCACATCATC CTCGAAAGCC
53581 ACCCCCGACC GCGCCCCGAA CCGGCGGCG CACCCGACAC CGGACCGCTG CCGCTGCTGC
53641 TCTCGGCCCC CACCCCGCAG GCACTCGACG CACAGGTACA CCGCCTGCGC GCGTCTCTCG
53701 ACGACAACCC CCGGCGCGGAC CCGGTGCGCG TCGCGCAGAC ACTCGCCCGG CGCACCCAGT
53761 TCGAGCACCG CCGCGTGCTG CTGCGCGACA CGTCACTCAC CGTGAGCCCG AACGCGGGCC
53821 GCGGACCGGT GGTCTTCGTC TACTCGGGGC AAAGCACGCT GCACCCGAC ACCGGGCGGC
53881 AACTCGGTC CACTACCCC GTGTTCGCG AAGCGTGGCG CGAGGCCCTC GACCACCTCG
53941 ACCCCACCCA GCGCCCGGCC AGCACTTTC CCGACAGAC CCGCTCACG CCGCTCTGCG
54001 GTCCTGGGG CATCACCCG CACGCTCTCA TCGGCGACTC CCTCGGTGAG ATCAACCGCG
54061 GCGACGCGCG GGTGCTCTTC TCCCTGAGCG AGCGGGGCGC GTCCTCACG ACCCGCACCG
54121 GCTGATGGA CCAACTGCC TCGGGCGCG CGATGGTCA CGTCTGACC AGCGAGGAAA
54181 AGGCACGCCA GGTGCTGCGC CCGGGCGTGG AGATCGCCGC CCGCAACGGC CCGACTCCG

SUBSTITUTE SHEET (RULE 26)

54241 TCGTGCTGTC CGGGGACGAG GAAGCCGTAC TCGAAGCCGC CCGGCAGCTC GGCATCCACC
54301 ACCGCTGTC GACCCGCCAC GCCGGCCACT CCGAGCGCAT GCAGCCACTC GTGCGCCGCC
54361 TCGTCGACGT CGCCCGGACC CTGACGTACC ACCAGCCCCA CACCGCCATC CCGGGCGACC
54421 TACCCACCGC CGAATACTGG GCGCACCGAG TCGCGACCA ACTACGTTTC CAGGGCGACA
54481 CCGAGCAGTA CCGGGGCGCG ACGTTCCCTG AGATCGGCCC CAACCAGGAC CTCTCGCCGC
54541 TCGTCGACGG CGTTCCCGCC CAGACCGGTA CCGCCGACGA GGTGCGGGCG CTGCACACCG
54601 CGCTCGCGCA GCTCCACGTC CGCGCGCTCG CGATCGACTG GACGCTCGTC CTCGCGGGGG
54661 ACCGCGCGCC CGTCACGCTG CCCACGTATC CGTTCCAGCA CAAGGACTAC TGGCTGCGGC
54721 CCACCTCCCG GGCCGATGTG ACCGGCGCGG GGCAGGAGCA GGTGGCGCAC CCGCTGCTCG
10 54781 GCGCCGCGST CGCGCTGCCC GGCACGGGCG GAGTCGTCCT GACCGGCCCG CTGTGCTGG
54841 CCTCCCATCC GTGGCTCGGC GAGCACGCGG TCGACGGCAC CGTGCTCCTG CCGGGCGCGG
54901 CCTTCCTCGA ACTCGCGGCG CGCGCCGGCG ACGAGGTCGG CTGCGACCTG CTGCACGAAC
54961 TCGTCATCGA GACGCCGCTC GTGCTGCCCC CGACCGGCGG TGTGGCGGTC TCGCTCGAGA
55021 TCGCCGAACC CGACGACAG CGCGGGCGGG CGGTACCGT CCACGCGCGG GCGGACGGT
15 55081 CGGGCCTGTG GACCCGACAC GCCGGCGGAT TCCTCGGCAC CAAGCCGCGG CCGGCCACGG
55141 CCACGGACCC GGCACCTGG CCGCCCGCGG AAGCCGGACC GGTGACGTC CCGGACGCT
55201 ACGACCGGTT CGAGGACATC GGGTACTCCT ACGGACCGGG CTTCGGGGG CTGCGGGCCG
55261 CCTGGCGCGC CGCGCACACC GTGTACGCCG AGGTGCGGCT CCGCGACGAG CAGAGCGCCG
55321 ACSCCGCCCG TTTCACGCTG CACCCCGCGC TGCTCGACGC CGCGTTCCAG GCGGGCGCGC
20 55381 TGSCCGCGCT CGACGCACCC GCGGGGGCGG CCGGACTGCC GTTCTCGTTC CAGGACGTCC
55441 GCATCCACGC CGCCCGGGCG ACGCGGTGCG GGTTCACGGT CGGCGCGCAC GCGGAGCGCA
55501 GCACCGTCCG CATGACCGGC CCGGACGGGC AGCTGGTGGC CGTGGTGGT CCGGTGCTGT
55561 CCGCGCCCTA CGCGGAAGGC TCCGGTGACG GCCTGCTGCG CCGGTCTGG ACCGAGCTGC
55621 CGATGCCCTT CCGTCCGCG GACGATCCGC GCGTGGAGGT CCTCGGCGCC GACCCGGGCG
25 55681 ACGGCGACGT TCCGGCGGCC ACCCGGGAGC TGACCGCCCG CGTCTCGGC GCGCTCCAGC
55741 GCCACCTGTC CGCCGCGGAG GACACCACCT TGGTGGTACG GACCGGCACC GGCCCGGCCG
55801 CTSCCGCCGC CGCGGGTCTG GTCCGCTCGG CGCAGGCGGA GAACCCCGCG CGCGTCGTGC
55861 TCGTCGAGGC GTCCCGGAC ACCTCGGTGG AGCTGCTCGC CGCGTGCGCC GCGCTGGACG
30 55921 AACCGCAGCT GGCCGTCCGG GACGGCGTGC TCTTCGCGCC CGCGTGGTC CGGATGTCCG
55981 ACCCCGCGCA CGGCCGCTG TCCCTGCCGG ACGGCGACTG GCTGCTCACC CGGTCCGCT
56041 CCGGCACGTT GCACGACGTC GCGCTCATAG CCGACGACAC GCGCCGCGG GCGCTCGAAG
56101 CCGGCGAGGT CCGCATCGAC GTCCGCGCGG CCGGACTGAA CTTCGCGAT GTGCTGATCG
56161 CGCTCGGGAC GTACACCGGG GCCACGGCCA TGGGCGGCGA GGCCGCGGGC GTCGTGGTGG
56221 AGACCGGGCC CGGCGTGGAC GACCTGTCCC CCGGCGACCG GGTGTTCCGG CTGACCCGGG
35 56281 GCGGCATCGG CCGGACGGCC GTCACCGACC GCGCGTGGCT GGCCCGGATC CCGGACGGCT
56341 GGAGCTTCAC CACGGCGGCG TCCGTCCCGA TCGTGTTCGC GACCGCGTGG TACGGCTGG
56401 TCGACCTCGG CACACTGCGC GCGGCGAGA AGGTCTCTGT CCACGCGGCC ACCGGCGGTC
56461 TCGGCATGGC CGCCGCACAG ATCGCCCGCC ACCTGGGCGC CGAGCTCTAC GCCACCGCCA
40 56521 GTACCGGCA GCAGCACGTC CTGCGCGCCG CCGGGCTGCC CGACACGCAC ATCGCGACT
56581 CTCGGACGAC CGCGTTCGG ACCGCTTTCC CGCGCATGGA CGTCGTCTTG AACGCGCTGA
56641 CCGCGAGTT CATCGACGCG TCGCTCGACC TGCTGGACGC CGACCGCCGG TTCGTGAGA
56701 TGGGCGGCAC CGAGCTGCGC GACCCGACCG CGATCGTCCC CGCTACCTG CCGTTCGACC
56761 TGCTGGACGC GGGCGCCGAC CGCATCGGCG AGATCCTGGG CGAACTGCTC CGGTGTTCG
45 56821 ACGCGGGCGC GCTGGAGCCG CTGCCGTTCC GTGCTGSSA CGTCCGGCAG GCACGCGACG
56881 CGCTCGGCTG GATGAGCCGC GCCCGCCACA TCGGCAAGAA CGTCTGACG CTGCCCCGGC
56941 CGCTCGACCC GGAGGCGGCC GTGCTCTCA CCGGCGGCTC CGGACGCTC GCGGCATCC
57001 TCGCCCGCCA CCTGCGCGAA CGGCTGTCT ACCTGCTGTC CCGGACGGCA CCGCCCGAGG
57061 GGACGCGCGG CGTCCACCTG CCCTGCGACG TCGGTGACCG GGACAGCTG GCGCGGGCCC
57121 TGGAGCGGGT GGACCGGCCG ATCACCAGCG TGGTGCACCT CGCCGGTGC CTGGACGACG
50 57181 GCACCGTCGC GTCGCTCACC CCGGAGCGTT TCGACACGGT GCTGCGCCCG AAGGCCGACG
57241 GCGCCTGGTA CCTGCACGAG CTGACGAAGG AGCAGGACCT CGCCGCGTTC GTGCTCTACT
57301 CGTCGGCCGC CGGCGTGCTC GGCAACGCCG GCCAGGGCAA CTACGTGCGC GCGAACGCGT
57361 TCGTCGACGC GCTCGCCGAG CTGCGCCACG GTTCCGGGCT GCGGGCCCTC TCCATCGGCT
57421 GGGGCTCTG GGAGGACGTG AGCGGGCTCA CCGCGGCGCT CCGCGAAGCC GACCGGGACC
55 57481 GGATCGGCGC CAGCGGTTTC CGGGCAATCA CCGCGCAACA GGGCATGCAC CTGTACGAGC
57541 CCGCCGCGCG CACCGGAAGT CCGGTGGTGG TCGCGGCGGC GCTCGACGAC GCGCGGACG
57601 TGCGCTGCTG GCGCGGCTG CGGCGGACGA CCGTCCGGCG GGCCGCGGTC CCGGAGTGT
57661 CCGCGGCGCA CCGGCTCGCC GCGCTGACCG GCGACGAGCT CCGCGAAGCG CTGCTGACCG
57721 TCGTCCGGGA GAGCACCGCC GCGGTGCTCG GCCACGTGGG TGGCGAGGAC ATCCCCGCGA
60 57781 CCGCGGCGCT CAAGGACCTC GGCATCGACT CGTCCACCGC GGTCCAGCTG CCGAACCGCC

57841 TCACCGAGGC GACCGGTGTG CGGCTGAACG CCACGGCGGT CTTCGACTTC CCGACCCGCG
 57901 ACGTGCTCGC CGGSAAGCTC GGCACGAAC TGACCGGCAC CCGCGCGCCC GTCTGCCCCC
 57961 GGACCGCGGC CACGGCCGGT GCGCACGACG AGCGGCTGGC GATCGTGGGA ATGGCCTGCG
 58021 GGCTGCCCCG CGGGSTCGCG TCACCGAGG AGCTGTGGCA CCTCGTGGCA TCCGGCAGCG
 58081 ACGCCATCAC GGAGTTCCCG ACGGACCGCG GCTGGGACST CGACGCGATC TACGACCCCG
 58141 ACCCCGACGC GATCGGCAAG ACCTTCGTCC GGCACGGTGG CTTCCTCACC GCGCGACACG
 58201 GCTTCGACGC GGCCTTCTTC GGCATCAGCC CCGCGGAGGC CCTCGCGATG GACCCGACGC
 58261 AGCGGGTGCT CCTGGAGACG TCGTGGGAGG CGTTCGAAAG CGCCGGCATC ACCCCGGACT
 58321 CGACCGCGCG CAGCGACACC GCGGTGTTCC TCGGCGCTTT CTCTACGGT TACGGCACCG
 58381 GTGCGGACAC CGACGGCTTC GCGCGGACCG GCTCGCAGAC CAGTGTGCTC TCCGGCCGGC
 58441 TGTCGTACTT CTACGGTCTG GAGGGTCCGG CGGTACGGT CGACACGGCG TGTTCGTCTG
 58501 CGCTGGTGGC GCTGCACCAG GCCGGGACGT CGCTGCGCTC CGGCGAATGC TCGCTCGCCC
 58561 TGGTCGGCGG CGTCACGGTG ATGGCGTCTC CCGGCGGCTT CGTGGAGTTC TCCCGGCAGC
 58621 GCGGCCTCGC GCCGGACGGC CGGCGGAAGG CGTTCGGCGC GGGTGCGGAC GGCACGAGCT
 58681 TCGCCGAGGG TGCCGGTGTG CTGATCGTCG AGAGGCTCTC CGACGCGGAA CGCAACGGTC
 58741 ACACCGTCTT GGCGGTCTCG CTGGTTCGG CGGTCAACCA GGATGGTGCC TCCAACGGGC
 58801 TGTCGGCGCC GAACGGGCGG GCGCGGACCG GGTGATCCG GCAGGCCCTG GCCAACGGCG
 58861 GGCTCACCCC GCGCGACGTG GACGCCGTCG AGGCCACCG CACCGGCACC AGGCTGGGCG
 58921 ACCCCATCGA GGCACAGGCG GTACTGGCCA CCTACGGACA GGAGCGCGCC ACCCCCCTGC
 58981 TGCTGGGCTC GCTGAAGTCC AACATCGGCC ACGCCAGGC CGCGTCCGGC GTCGCCGGCA
 59041 TCATCAAGAT GGTGCAGGCC CTCCGGCACG GGGAGCTGCC GCCGACGCTG CACGCCGACG
 59101 AGCCGTCCGC GCACGTCGAC TGGACGGCGG GCGCCGTGCA ACTGCTGACG TCGGCCCGGC
 59161 CGTGGCCCGA GACCGACCG CCACGGCGTG CCGCGTCTC CTCGTTCGGG GTGAGCGGCA
 59221 CCAACGCCCC CGTCATCCTG GAGGCGGAC CGGTAACGGA GACGCCCGCG GCATCGCCTT
 59281 CCGGTGACCT TCCCTGCTG GTGTGCGCAC GCTCACCGGA AGCGCTCGAC GAGCAGATCC
 59341 GCCGACTGCG CGCCTACCTG GACACCACCC CGGACGTGCA CCGGGTGGCC GTGGCACAGA
 59401 CGCTGGCCCC GCGCACACAC TTCGCCCCACC GCGCCGTGCT GCTCGGTGAC ACCGTCATCA
 59461 CCACACCCCC CGCGGACCGG CCCGACGAAC TCGTCTTCGT CTACTCCGGC CAGGGCACCC
 59521 AGCATCCCCG GATGGGCGAG CAGCTCGCCG CCGCCCATCC CGTGTTCGCG GACGCTGGC
 59581 ATGAAGCGCT CCGCGCCTT GACAACCCCG ACCCCACAGA CCCCACGAC AGCCAGCATG
 59641 TGCTCTTCGC CCACCAGGCG GCGTTCACCG CCCTCCTGCG GTCCTGGGGC ATCACCCTGC
 59701 ACGCGGTCAT CGGCCACTCG CTGGGCGAGA TCACCGCGGC GCACGCGGCC GGCATCTGT
 59761 CGCTGGACGA CGCGTGCACC CTGATCACC ACGCGCGCCG CCTCATGCAC ACGTCCCGC
 59821 CACCCGGTGC CATGGTCACC GTACTGACCA GCGAAGAGAA GGCACGCCAG GCGTTGCGGC
 59881 CGGGCGTGGA GATCGCCGCC GTCAACGGGG CCCACTCCAT CGTGTGTGCC GGGGACGAGG
 59941 ACGCCGTGCT CACCGTCGCC GGGCAGTCTG GCATCCACCA CCGCTGCCAC GCGCCGACG
 60001 CCGGGCACATC CGCGCACATG GCGCCGTGG CCGCCGAGCT GCTCGCCACC ACCCGCGGGC
 60061 TCCGCTACCA CCTCCCCAC ACCTCCATTC CGAACGACCC CACCACCGCT GAGTACTGGG
 60121 CCGAGCAGGT CCGCAAGCCC GTGCTGTTCC ACGCCACCG GCAGCAGTAC CCGGACGCCG
 60181 TCTTCGTGSA GATCGGCCCC GCGCAGGACC TCTCCCCCT CGTCGACGGG ATCCCGCTGC
 60241 AGAACGGCAC CGCGGACGAG GTGCACGCGC TGCACACCGC GCTCGCGCAC CTCTACGCGC
 60301 CCGGTGCCAC GCTCGACTGG CCGCGCATCC TCGGGGCTGG GTCACGGCAC GACGCGGATG
 60361 TGCCCGCGTA CGCGTTCCAA CGCGGCACT ACTGGATCGA GTCGCGACG CCGGCCGAT
 60421 CCGACGCGGG CCACCCCGTG CTGGGCTCCG GTATCGCCCT CGCCGGGTG CCGGGCGGG
 60481 TGTTACAGGG TTCTGTGCCG ACCGGTGGCG ACCGCGCGGT GTTCGTGCGC GAGCTGGCGC
 60541 TGGCCCGCGC GGACGCGGTC GACTGCGCCA CCGTCGAGCG GCTCGACATC CCTCCGTGC
 60601 CCGGCCCGCC GGGCCATGGC CGGACGACCG TACAGACCTG GGTGACGAG CCGGCGGACG
 60661 ACGGCCGGCG CCGGTTACCC GTGCACACCC GCACCGGCGA CGCCCGCTGG ACGCTGCACG
 60721 CCGAGGGGGT GCTGCGCCCC CATGGCACGG CCTGCCCCGA TCGGGCGCAC GCGGAGTGGC
 60781 CCCCACCGGG CGCGGTGCC CCGGACGGGC TGCCGGGTGT GTGGCGCCG GGGGACGAGG
 60841 TCTTCGCCGA GGCCGAGGTG GACGGACCGG ACGGTTTCGT GGTGCACCCC GACCTGCTCG
 60901 ACGCGGTCTT CTCCGCGGTC GCGGACGGAA GCCGCCAGCC GGCCGGATGG CGCGACCTGA
 60961 CCGTGACGCG GTCGGACGCC ACCGTACTGC GCGCTGCCT CACCCGCGCG ACCGACGGAG
 61021 CCATGGGATT CGCCGCTTC GACGGCGCCG GCCTGCCGGT ACTACCGCG GAGGCGGTGA
 61081 CGCTGCGGGA GGTGGCGTCA CCGTCCGGCT CCGAGGAGTC GGACGGCCTG CACCGGTTGG
 61141 AGTGGCTCGC GGTGCGCGAG GCGGTCTACG ACGGTGACCT GCGCGAGGGA CATGTCCTGA
 61201 TCACCGCGCT CGACCCCGAC GACCCCGAGG ACATACCCAC CCGCGCCAC ACCCGCGCGA
 61261 CCGCGCTCTT GACCGCCCTG CAACACCAAC TCACCAACAC CGACCAACAC CTCATCGTCC
 61321 ACACCAACAC CGACCCCGCC GCGCGCACCG TCACCGGCTT CACCCGACCC GCGCAGAACG
 61381 AACACCCCAA CCGCATCCGC CTCATCGAAA CCGACCAACC CCACACCCCC CTCCCCCTGG

SUBSTITUTE SHEET (RULE 26)

	61441	CCCAACTCGC	CACCCTCGAC	CACCCCAACC	TCCGCCTCAC	CCACCACACC	CTCCACCACC
	61501	CCCACCTCAC	CCCCCTCCAC	ACCACCACCC	CACCCACCAC	CACCCCCCTC	AACCCCGAAC
	61561	ACGCCATCAT	CATCACCAGC	GGCTCCGGCA	CCCTCGCCGG	CATCCTCGCC	CGCCACCTGA
5	61621	ACCAACCCCA	CACCTACCTC	CTCTCCCGCA	CCCCACCCCC	CGACGCCACC	CCCGGCAACC
	61681	ACCTCCCTCG	CGACGTCGGC	GACCCCAACC	AACCTCGCCAC	CACCCCTCACC	CACATCCCCC
	61741	AACCCCTCAC	CGCCATCTTC	CACACCAGCG	CCACCCTCGA	CGACGGCATC	CTCCACGCCC
	61801	TCACCCCGCA	CGCCCTCACC	ACCGTCTCTC	ACCCCAAAAGC	CAACGCCGCC	TGGCACCTGC
	61861	ACCACCTCAC	CCAAAACCAA	CTCCCTCACCC	ACTTCGTCTT	CTACTCCAGC	GCCGCCGCCG
	61921	TCCTCGGCAG	CCCCGGACAA	GGAAACTACG	CCGCCGCCAA	CGCCTTCTCT	GACGCCCTCG
10	61981	CCACCCACCG	CCACACCCTC	GGCCCAACCC	CCACCTCCAT	CGCCTGGGGC	ATGTGGCACA
	62041	CCACCAGCAC	CCTCACCAGG	CAACTCGACG	ACGCCGACCG	GGACCGCATC	CGCCGCGGCG
	62101	GTTTCTCTCC	GATCACCAGG	GACGAGGGCA	TGCGCCTCTA	CGAGGCGGCC	GTCGGCTCCG
	62161	GCGAGGACTT	CGTCATGGCC	GCCGCGATGG	ACCCGGCACA	GCCGATGACC	GGCTCCGTAC
15	62221	CGCCATCTCT	GAGCGGCTGT	CGCAGGAGCG	CGCGGCGCGT	CGCCCGTGCC	GGGCAGACGT
	62281	TCGCCACGCG	GCTCGCCGAG	CTGCCGACG	CCGACCGCGG	CGCGGCGCTG	ACCACCTCTG
	62341	TCTCGGACGC	CACGGCCGCC	GTGCTCGGCC	ACGCCGACGC	CTCCGAGATC	CGCCCGACCA
	62401	CGACGTTCAA	GGACCTCGGC	ATCGACTCGC	TCACCGCGAT	CGAGCTGCGC	AACCGGCTCG
	62461	CGGAGGCGAC	CGGGCTGCGG	CTGAGTGCCA	CGCTGGTGTT	CGACCACCCG	ACACCTCGGG
20	62521	TCCTCGCCGC	CAAGCTCCGC	ACCGATCTGT	TCGGCACGGC	CGTGCCACAG	CCCGCGCGGA
	62581	CGGCACGGAC	CCACCACGAC	GAGCCACTCG	CGATCGTCGG	CATGGCGTGC	CGACTGCCCG
	62641	CGGGGTCGCG	CTCGCCGGAG	GACCTTGTCG	AGCTCGTGGC	GTCCGGCACC	GACGCGATCA
	62701	CCGAGTTCCC	CACCGACCGC	GGCTGGGACA	TCGACCGGCT	GTTCGACCCG	GACCCGGACG
	62761	CCCCCGGCAA	GACCTACGTC	CGGCACGGCG	GCTTCTCTCG	CGAGGCCGCC	GGCTTCGATG
25	62821	CCGCGTTCTT	CGGCATCAGC	CCGCGCGAGG	CACGGGCCAT	GGACCCGCGC	CAGCGCGTCA
	62881	TCCTCGAAAC	CTCCTGGGAG	GCGTTTCGAG	ACGCGGGCAT	CGTGCCGGAC	ACGCTGCGCG
	62941	GCAGCGACAC	CGGCGTGTTT	ATGGGCGCGT	TCTCCCATGG	GTACGGCGCC	GGCGTCGACC
	63001	TGGGCGGGTT	CGGCGCCACC	GCCACGCAGA	ACAGCGTGCT	CTCCGGCCCG	TTGTCTGACT
	63061	TCTTCGGCAT	GGAGGGCCCG	GCCGTACCCG	TCGACACCCG	CTGCTCGTGC	TCGCTGGTCC
30	63121	CCCTGCACCA	GGCGGCACAG	GCGCTGCGGA	CTGGAGAATG	CTCGCTGGCG	CTCGCCGGCG
	63181	GTGTACAGGT	GATGCCACAC	CCGCTGGGCT	ACGTGAGATT	CTGCCGCCAG	CGGGGACTCG
	63241	CCCCCGACGG	CCGTTGCCAG	GCCTTCGCGG	AAGGCGCCGA	CGGCACGAGC	TTCTCGGAGG
	63301	GCGCCGGCGT	TCTTGTGCTG	GAGCGGCTCT	CCGACGCCGA	GCGCAACGGA	CACACCGTCC
	63361	TCGCGGTCTG	CCGCTCCTCC	GCCGTCAACC	AGGACGGCGC	CTCCAACGGC	ATCTCCGCAC
35	63421	CCAACGGCCC	CTCCCAGCAG	CGCGTCAATC	GCCAGGCCCT	CGACAAGGCC	GGGCTCGCCC
	63481	CCGCCGACGT	GGACGTGGTG	GAGGCCACCG	GCACCGGAAC	CCCGCTGGGC	GACCCGATCG
	63541	AGGCACAGGC	CATCATCGCG	ACCTACGGCC	AGGACCGCGA	CACACCGCTC	TACCTCGGTT
	63601	CGGTCAAGTC	GAACATCGGA	CACACCCAGA	CCACCGCCGG	TGTCGCCGGC	GTCATCAAGA
	63661	TGGTCATGGC	GATGCGCCAC	GGCATCGCGC	CGAAGACACT	GCACGTGGAC	GAGCCGTCGT
40	63721	CGCATGTGGA	CTGGACCGAG	GGTGCGGTGG	AACGTGCTAC	CGAGGCGAGG	CCGTGGCCCC
	63781	ACGCGGGACG	CCCGCGCCGC	GCGGGCGTGT	CGTCGCTCGG	TATCAGCGGT	ACGAACGCCC
	63841	ACGTGATCCT	TGAGGGTGTT	CCCGGGCCGT	CGCGTGTGGA	GCCGTCTGTT	GACGGGTTGG
	63901	TGCCGTTGCC	GTTGTGCGCT	CGAGTGTAGG	CGAGTCTGCG	GGGGCAGGTG	GAGCGGCTAG
	63961	AGGGGTATCT	GCGCGGGAGT	GTGGATGTGG	CCGCGGTTCG	GTCGCTGAGC	GTGCTGAGG
45	64021	GTGCTGTCTT	CGGTCACCGT	GCGGTACTGC	TGGGTGATGC	CCGGGTGATG	GGTGTGGCGG
	64081	TGGATCAGCC	GCGTACGGTG	TTCGTCTTTC	CCGGGCAGGG	TGCTCAGTGG	GTGGGCATGG
	64141	GTGTGGAGTT	GATGGACCGT	TCTGCGGTGT	TCGCGGCTCG	TATGGAGGAG	TGTGCGCGGG
	64201	CGTTGTTGCC	CCACACGGGC	TGGGATGTGC	GGGAGATGTT	GGCGCGGCCG	GATGTGGCGG
50	64261	AGCGGGTGGA	GTTGGTCCAG	CCGGCCAGCT	GGGCGTTCGC	GGTCAGCCTG	GCCGCACTGT
	64321	GGCAGGCCCA	CGGGGTCTGA	CCCGACGCGG	TGATCGGACA	CTCCAGGGG	GAGATCGCGG
	64381	CGGCGTGCGT	GCGCGGGGCC	CTCAGCCTTG	AGGACGCCGC	CCGCGTGGTG	GCCTTGCGCA
	64441	GCCAGGTCAT	CGCGGCGCGA	CTGGCCGGGC	GGGGAGCGAT	GGCTTCGGTG	GCATTGCCGG
	64501	CCGGTGAGGT	CGGTCTGGTC	GAGGGCGTGT	GGATCGCGGC	GCGTAACGGC	CCCGCCTCGA
	64561	CAGTCGTGGC	CGGCGAGCCG	TCGGCGGTGG	AGGACGTGGT	GACGCGGTAT	GAGACCGAAG
55	64621	GCGTGCGAGT	GCGTCGTATC	GCCGTCGACT	ACGCCTCCCA	CACGCCCCAC	GTGGAAGCCA
	64681	TCGAGGACGA	ACTCGCTAGG	GTAAGGAAGG	CAGTTGCAGG	GAAGGCCGCG	TCGGTGGCGT
	64741	GTTGGTCGAC	CGTGACAGAG	CCCTGGGTGA	CCGAGCCGGT	GGATGAGAGT	TACTGGTACC
	64801	GGAACCTGCG	TGCCCCCGTC	GCGCTGGACG	CGGCGGTGGC	GGAGCTGGAC	GGGTCCGTGT
	64861	TCGTGGAGTG	CAGCGCCCAT	CCGGTGCTGC	TGCCGCGCAT	GGAACAGGCC	CACACGGTGG
60	64921	CGTCGTTGCG	CACCGGTGAC	GGCGGCTGGG	AGCGATGGCT	GACGGCGTTG	GCGCAGGCGT
	64981	GGACCCTGGG	GCGGGCAGTG	GACTGGGACA	CGGTGGTCTGA	ACCGGTGCCA	GGGCGGCTGC

SUBSTITUTE SHEET (RULE 26)

	65041	TCGATCTGCC	CACCTACGCG	TTCGAGCGCC	GGCGCTACTG	GCTGGAAGCG	GCCGGTGCCA
	65101	CCGACCTGTC	CGCGGCCGGG	CTGACAGGGG	CAGCACATCC	CATGCTGGCC	GCCATCACGG
	65161	CACTACCCGC	CGACGACGGT	GGTGTGTGTC	TCACCGGCCG	GATCTCGTTG	CGCACGCATC
5	65221	TCTGGCTGGC	TGATCACGCG	GTGCGGGGCA	CGGTCTGTCT	GCCGGGCACG	GCCCTTGTGG
	65281	AGCTGGTCAT	CCGGGCCGGT	GACGAGACCG	GTTCGGGGAT	AGTGGATGAA	CTGGTCATCG
	65341	ATCCCCCCT	CGTGGTGCCG	GCGACCGCAG	CCGTGGATCT	GTCCGTGACC	GTGGAAGGAG
	65401	CTGACGAGGC	CGGACGGCGG	CGAGTGACCG	TCCACGCCCC	CACCGAAGGC	ACCGGCAGCT
	65461	CGACCCGGCA	CGCCAGCGGC	ACCCTGACCC	CCGACACCCC	CGACACCCCC	AACGCTTCCG
10	65521	GTGTTGTCCG	TGCGGAGCCG	TTCTCGCAGT	GGCCACCTGC	CACTGCCCGC	GCCCTCGACA
	65581	CCTCCGAGTT	CTACTTGC GC	CTGGACGC GC	TGGGCTACCG	GTTTCGACCC	ATGTTCCGCG
	65641	GAATGCGGGC	TGCCTGGCGT	GATGGTGACA	CCGTGTACGC	CGAGGTCGCG	CTCCCGGAGG
	65701	ACCGTGCCGC	CGACGCGGAC	GGTTTCGGCA	TGCACCCGGC	GCTGCTCGAC	GCGGCCTTGC
	65761	AGAGCGGCAG	CCTGCTCATG	CTGGAATCGG	ACGGCGAGCA	GAGCGTGCAA	CTGCCGTTCT
	65821	CCTGGCACGG	CGTCCGGTTC	CACGCGACGG	GCGCGACCAT	GCTGCGGGTG	GCGGTCGTAC
15	65881	CGGGCCCGGA	CGGCCTCCGG	CTGCATGCCG	CGGACAGCGG	GAACCGTCCC	GTCCGCGACGA
	65941	TCGACGCGCT	CGTGACCCGG	TCCCCGGAAG	CGGACCTCGC	GCCCGCCGAT	CCGATGCTGC
	66001	GGGTGCGGGT	GGCCCCGGTG	CCCGTACCTG	CCGGGGCCGG	TCCGTCCGAC	CCGGACCTGC
	66061	TGACGCTGCG	CGGCGACGAC	GCCGACCCGC	TCGGGGAGAC	CCGGGACCTG	ACCACCCGTG
20	66121	TTCTCGACGC	GCTGCTCCGG	GCCGACCGGC	CGGTGATCTT	CCAGGTGACC	GGTGGCCTCG
	66181	CCGCCAAGGC	GGCCGCGAGC	CTGGTCCGCA	CCGCTCAGAA	CGAGCAGCCC	GGCCGCTTCT
	66241	TCCTCGTCGA	AACGGACCCG	GGAGAGGTCC	TGGACGGCGC	GAAGCGCGAC	GCGATCGCGG
	66301	GACGCGGCA	GCCCCATGTG	CGGCTGCGCG	ACGGCCTCTT	CGAGGCAGCC	CGGCTGATGC
	66361	GGGCCACGCC	GTCCCTGACG	CTCCCGGACA	CCGGGTCTGT	GCAGCTGCGG	CCGTCCGCCA
25	66421	CCGGTTCCCT	CGACGACCTT	GCCGTCGTCC	CCACCGACGC	CCCGGACCCG	CCGTCCGCGG
	66481	CCGGCGAGGT	GCGGATCGCG	GTACGCGCGG	CGGGCCTGAA	CTTCCGGGAT	GTCACGGTGC
	66541	CGCTCGGTGT	GGTCGCGGAT	GCGCGTCCGC	TCGGCAGCGA	GGCCGCGGGT	GTCGTCTTGG
	66601	AGACCGGCCC	CGGTGTGCAC	GACCTGGCGC	CCGGCGACCG	GGTCTGGGGG	ATGCTCGCGG
	66661	GCSCCTTCGG	ACCGGTTCGG	ATCACCGACC	GGCGGCTGCT	CGGCCGGATG	CCGGACGGCT
30	66721	GACAGTTCCT	GCAGGCGGCG	TCCGTGATGA	CCCGGTTTCG	GACCGCGTGG	TACGSCCTGG
	66781	TCGACCTGGC	CGGGCTGCGC	CCCGGCGAGA	AGGTCTTGAT	CCACGCGCGG	GCGACCGGTG
	66841	TCGGCGCGGC	GGCCGTCCAG	ATCGCGCGGC	ATCTGGGCGC	GGAGGTGTAC	GCGACCAACA
	66901	GCGCCGCGAA	GCGCCATCTG	GTGGACCTGG	ACGGAGCGCA	TCTGGCCGAT	TCCCGCAGCA
	66961	CCGCGTTCGC	CGACGCGTTC	CCGCGCGTCC	ATGTCGTGCT	CAACTCGCTC	ACCGGTGAAT
35	67021	TCCTCGACGC	GTCCGTTCGG	CTGCTCGCGG	CGGGTGGCCG	GTTTCATCGAG	ATGGGGAAGA
	67081	CGGACATCCG	GCACGCCGTC	CAGCAGCCGT	TGACCTTGAT	GGACGCGCGG	CCCGACCGGA
	67141	TGCAGCGGAT	CATCGTCGAG	CTGCTCGGCC	TGTTTCGCGG	CGACGTGCTG	CACCGCTTGC
	67201	CGGTCCACGC	CTGGGACGTG	CGGCAGGCGC	GGGAGGCGTT	CGGCTGGATG	AGCAACGGGC
	67261	GTCACACCGG	CAAGCTGGTG	CTGACGGTCC	CGCGGCCGCT	GGATCCCGAG	GGGGCCGTCG
40	67321	TCATCACCGG	CGGCTCCGGC	ACCCTCGCCG	GCATCCTCGC	CCGCCACCTG	GGCCACCCCC
	67381	ACACCTACCT	GCTCTCCCGC	ACCCACCCCC	CCGACACCAC	CCCCGGCACC	CACCTCCCTT
	67441	GCGACGTCGG	CGACCCCCAC	CAACTCGCCA	CCACCCTCGC	CCGCATCCCC	CAACCCCTCA
45	67501	CCGCCGTCTT	CCACACCGCC	GGAACCTTCG	ACGACGCCCT	GCTCGACAAC	CTCACCCCCG
	67561	ACCGCGTCGA	CACCGTCTTC	AAACCCAAGG	CCGACGCCGC	CTGGCACCTG	CACCGCTCA
	67621	CCCGCGACAC	CGACCTCGCC	GCGTTCGTGC	TCTACTCCGC	GGTCGCCGGC	CTCATGGGCA
	67681	GCCCGGGGCA	GGGCAACTAC	GTGCGGGCGA	ACGCGTTTCT	CGACGCGCTC	GCCGAACACC
	67741	GCCGTGCGCA	AGGGCTGCCC	GCGCAGTCCC	TCGCATGGGG	CATGTGGGCG	GACGTCAGCG
	67801	CGCTCACCGC	GAAACTCACC	GACGCGGACC	GCCAGCGCAT	CCGGCGCAGC	GGATTCCCGC
	67861	CSTTGAGCGC	CGCGGACGGC	ATGCGGCTGT	TCGACGCGGC	GACGCGTACC	CCGGAACCGG
50	67921	TCGTCTGTCG	GACGACCGTC	GACCTCACCC	AGCTCGACGG	CGCCGTCGCG	CCCTTGCTCC
	67981	GCGGTCTGGC	CGCGCACCGG	GCCGGGCCGG	CGCGCACGGT	CGCCCGCAAC	GCCGCGGAAG
	68041	AGCCCTTGGC	CGTGCGTCTT	GCCGGGCGTA	CCGCCGCCGA	GCAGCGGCGC	ATCATGCGAG
	68101	AGGTCTGTCT	CCGCCACGCG	GCCGCGGTCC	TCGCGTACGG	GCTGGGCGAC	CGCGTGCGCG
	68161	CGGACCGTCC	GTTCCGCGAG	CTCGGTTTTC	ATTTCGTGAC	CGCGGTGAC	CTGCGCAATC
55	68221	GGCTCGCGGC	CGAGACGGGG	CTGCGGCTGC	CGACGACGCT	GGTGTTCAGC	CACCCGACGG
	68281	CGGAGGCGCT	CACCGCCAC	CTGCTCGACC	TGATCGACGC	TCCACCGGCC	CGGATCGCCG
	68341	GGGAGTCCCT	GCCCGCGGTG	ACGGCCGCTC	CCGTGGCGGC	CGCGCGGAC	GCGACGAGC
	68401	CGATCGCCAT	CGTGCGGATG	GCGTGCGGCT	TGTCGCGTGG	TGTGACGTCG	CCCGAGGACC
	68461	TGTGGCGGCT	CGTCGAGTCC	GCGACCGACG	CGATCACCAC	GCCTCCTGAC	GACCGCGGCT
60	68521	GGGACGTCGA	CGCGCTGTAC	GACGCGGACC	CGGACGCGGC	CGGCAAGGCG	TACACCTGCG
	68581	GGGGCGGTTA	CCTGGCCGGG	GCGGCGGAGT	TCGACGCGGC	GTTCTTCGAC	ATCATCTCCG

SUBSTITUTE SHEET (RULE 26)

68641 GCGAAGCGCT CGGCATGGAC CCGCAGCAAC GCCTGCTGCT CGAAACGGCG TGGGAGGGGA
 68701 TCGAGCGCGG CCGGATCAGT CCGGCGTCGC TCCGCGGCGG GGAGGTGCGC GTCTATGTGG
 68761 GTGCGGCCGC GCAGGGCTAC GGGCTGGGCG CCGAGGACAC CGAGGGCCAC GCGATCACCG
 5 68821 GTGGTTCCAC GAGCCTGCTG TCCGGACGGC TGGCGTACGT GCTCGGGCTG GAGGGCCCGG
 68881 CCGTCACCGT GGACACGGCG TGCTCGTCGT CTCTGGTCGC GCTGCATCTG GCGTGCCAGG
 68941 GGCTGCGCCT GGGCGAGTGC GAACTCGCTC TGGCCGGAGG GGTCTCCGTA CTGAGTTCCG
 69001 CGGCCGCGTT CGTGGAGTTC TCCCGCCAGC GCGGGCTGCG GGCCGACGGG CGCTGCAAGT
 69061 CGTTCGGCGC GGGCGCGGAC GGCACGACGT AGTCCGAGGG CGTGGGCGTG CTCGTACTGG
 10 69121 AACGGCTCTC CGACGCCGAG CCGCTCGGGC ACACCGTGCT CGCCGTCGTC CGCGGCAGCG
 69181 CCGTCACGTC CGACGCCGCC TCCAACGGCC TCACCGCGCC GAACGGGCTC TCGCAGCAGC
 69241 GGGTCATCCG GAAGGCCTC GCGCGGCGCG GGCTGACCGG CGCCGACGTG GACGTCGTGG
 69301 AGGGGCACGG CACCGGCACC CCGCTCGGCG ACCCGGTGSA GCGGACGCG CTGCTCGCGA
 69361 GTACGGGCA GGACCGTCCG GCACCGGTCT GGCTGGGCTC GCTGAAGTCG AACATCGGAC
 69421 ATGCCACGGC CGCGGCCGCT GTCGCGGGCG TCATCAAGAT GGTGCAGGCG ATCGGCGCGG
 15 69481 GCACGATGCC GCGGACGCTG CATGTGGAGG AGCCCTCGCC CGCGGTCGAC TGGAGCACCG
 69541 GACAGGTGTC CCTGCTCGGC TCCAACGGCG CCTGGCCGGA CGACGAGCGT CCGCGCCGGG
 69601 CGGCCGTCTC CGCGTTCGGG CTCAGCGGGA CGAACGCGCA CGTCATCCTG GAACAGCACC
 69661 GTCCGCGCGC CGTGGCGTCC CAGCCGCCCG GCGCGCCCGG TGAGGAGTCC CAGCCGCTGC
 69721 CGTGGGTGCT CTCCGCGCGG ACTCCGGCCG CGCTGCGGGC CCAGGCGGCC CCGCTGCGCG
 20 69781 ACCACCTCGG GCGGCGACCG GACGCGGATC CGTTGGACAT CCGGTACCGG CTGGCCACCA
 69841 GTCCGCCCCA GTTCGCCAC CGTGCCCGG TCGTCGCCAC CACCCCGGAC GGATTCCGTG
 69901 CCGCGCTCGA CCGCCTCGCG GACGGCGCGG AGTCCGCCCG AGTCGTACCG GGGACCGCTC
 69961 AGGAGCGGCG CGTCGCCTTC CTCTTCGACG GCCAGGCGCG CCAGCGCGCC GGAATGGGGC
 70021 GCGAGCTCCA CCGCCGTTTC CCCGTCTTCG CCGCCGCGTG GGACGAGGTC TCCGACGCGT
 25 70081 TCGGCAAGCA CCTCAAGCAC TCCCCACGG ACGTCTACCA CCGCGAACAC GCGCTCTCG
 70141 CCCATGACAC CCTGTACGCC CAGGCCGGCC TGTTACAGCT CGAAGTGGCG CTGCTGCGCG
 70201 TGCTGGAGCA CTGGGGGGTG CCGCCGGACG TGCTCGTCGG GCACTCCGTC GCGGAGGTGA
 70261 CCGCGCGCGTA CGCGCGGGG GTGCTCACCC TGGCGGACCG GACGGAGTTG ATCGTGGCCC
 30 70321 GGGGGCGGGC GCTGCGGGCG CTGCGCCCGG GGGCGATGCT CGCCGTCGAC CGCCGCCCGG
 70381 CCGAGGTGCG CGCCCGCACG GATCTGGACA TCGCCGCGGT CAACGGCCCCG TCCGCCGTGG
 70441 TGCTCGCCCG TTCGCCGAC GATGTGGCGG CGTTTGAACG GGAGTGGTGG GCGGCGGGCG
 70501 GCGCGACGAA ACGGCTCGAC GTCGGGCACG CGTTCCACTC CCGGCACGTC GACGGTGGCG
 70561 TCGACGGCTT CCGTACGGTG CTGGAGTCGG TCGCGTTCGG CGCGGCGCGG CTGCCGGTGG
 35 70621 TGTCCACGAC GACGGGCCGG GACGCGCGG ACGACCTCAT AACGCCCGCG CACTGGCTGC
 70681 GCCATGCGCG TCGGCGGGTG CTGTTCTCGG ATGCCGTCCG GGAGCTGGCC GACCGCGCGG
 70741 TCACCACGTT CGTGGCCGTC GGCCCTCCG GCTCCCTGGC GTCCGCGCGG GCGGAGAGCG
 70801 CCGGGGAGGA CGCCGGGACC TACCACGCGG TGCTGCGCGC CCGGACCGGT GAGGAGACCG
 70861 CCGCGCTGAC CGCCCTCGCC GAGCTGCACG CCCACGGCGT CCGGTCGAC CTGGCCGCGG
 70921 TACTGGCCGG TGGCCGGCCA GTGGACCTTC CCGTGTACGC GTTCCAGCAC CGTTCTACT
 40 70981 GGCTGGCCCC GGCCGTGGCG GGGGCGCCGG CCACCGTGGC GGACACCGGG GGTCCGGCGG
 71041 AGTCCGAGCC GGAGGACCTC ACCGTGCGCG AGATCGTCCG TCGGCGCACC GCGGCGCTGC
 71101 TCGGCGTAC GGACCCCGCC GACGTCGATG CGGAAGCGAC GTTCTTCGCG CTCGGTTTCG
 71161 ACTACTGGC GGTGCAGCGG CTGCGCAACC AGCTCGCCTC GGCAACCGGG CTGGACCTGC
 71221 CCGCGGCGGT CCTGTTTCGAC CACGACACCC CCGCCGCGCT CACCGCGTTC CTCCAGGACC
 45 71281 GGATCGAGGC CCGCCAGGAC CGGATCGAGG CCGGCGAGGA CGACGACGCG CCCACCGTGC
 71341 TCTCGCTCCT GGAGGAGATG GAGTCGCTCG ACGCCGCGGA CATCGCGGCG ACGCCGGCCC
 71401 CCGAGCGTGC GGCCATCGCC GATCTGCTCG ACAAGCTCGC CCATACCTGG AAGGACTACC
 71461 GATGAGCAC GATACGCACG AGGGAACGCG GCCCGCCGGC CGCTGCCCAT TCGCATCCA
 50 71521 GGACGGTAC CGCGCCATCC TGGAGAGCGG CACGGTGGGT TCGTTCGACC TGTTCGGCGT
 71581 CAAGCACTGG CTGGTTCGCG CCGCCGAGGA CGTCAAGCTG GTCACCAACG ATCCGCGGTT
 71641 CAGCTCGGCC GCGCCCTCCG AGATGCTGCC CGACCGGCGG CCGGGCTGGT TCTCCGGGAT
 71701 GGAATCACCG GAGCACAACC GCTACCGCCA GAAGATCCCG GGGGACTTCA CACTGCGCGC
 71761 GCGCGCAAG CCGGAGGACT TCGTCCCGGA GGCCGCGGAC GCCTGCTGG ACGACATCGA
 71821 GCGCGCGGGA CCGGCGACCG ACCTCATCCC CCGGTACGCC AAGCGGCTGC CCTCCCTCGT
 55 71881 CATCAACGCG CTGTACGGG TCACCCCTGA GGAGGGGGCC GTGCTGGAG CACGGATCGG
 71941 CGACATCACG GGCTCGGCCG ATCTGCTAG CGTCAAGACG CTGACGACG ACTTCTTCGG
 72001 GCACGCGCTG CCGCTGCTCC GCGCGAAGCG TGACGAGCGG GCGGAGGACC TGCTGCACCG
 72061 GCTGGCCTCG GCGACGACG GCGAGATCTC GCTCAGCGAC GACGAGGCGA CCGGCGTGT
 72121 CCGGACGCTG CTGTTTCGCG GCCACGACTC GGTGCAGCAG ATGGTTCGGT ACTGCTCTA
 60 72181 CGCACTGCTC AGCCACCCCG AGCAGCAGGC GCGCTGCGC GCGCGCCCGG AGCTGGTCCA

SUBSTITUTE SHEET (RULE 26)

72241 CAACGCGGTC GAGGAGATGC TCCGTTTCCT GCCCGTCAAC CAGATGGGCG TACCGCGCGT
 72301 CTGTGTCGAG GACGTCGATG TCGGGGGCGT GCGCATCCGT GCGGGGCGACA ACGTGATCCC
 72361 GCTCTACTCG ACGGCCAACC GCGACCCCGA GGTGTTCCCG CAGCCCGACA CCTTCGATGT
 72421 GACGCGCCCG CTGGAGGGCA ACTTCGCGTT CGGCCACGGC ATTCACAAGT GTCCCGGCCA
 5 72481 GCACATCGCC CGGGTGCTCA TCAAGGTGCG CTGCCTGCGG TTGTTGAGC GTTTCCCGGA
 72541 CGTCCGGCTG GCCGGCGACG TGCCGATGAA CGAGGGGCTC GGGCTGTTCA GCCCGGCCGA
 72601 GCTGCGGGTC ACCTGGGGGG CGGCATGAGT CACCCGGTGG AGACGTTGCG GTTGCCGAAC
 72661 GGGAGCACGG TCGCGCACAT CAACGCGGGG GAGGCGCAGT TCCTCTACCG GGAGATCTTC
 10 72721 ACCCAGCGCT GCTACCTGCG CCACGGTGTC GACCTGCGCC CGGGGGACGT GGTGTTGAC
 72781 GTCGGCGCGA ACATCGGCAT GTTCACGCTT TTCGCGCATC TGGAGTGTC TGGTGTGACC
 72841 GTGCACGCCT TCGAGCCCGC GCCCGTGCCG TTCGCGGCGC TCGGGGCGAA CGTACGCGG
 72901 CACGGCATCC CGGGCCAGGC GGACCAAGTG GCGGTCTCCG ACAGCTCCGG CACCCGGAAG
 72961 ATGACCTTCT ATCCCGACGC CACGCTGATG TCCGGTTTCC ACGGCGATGC CGCGGCCCGG
 73021 ACGGAGCTGT TGCGCACGCT CGGCCTCAAC GGCGGCTACA CCGCCGAGGA CGTCGACACC
 15 73081 ATGCTCGCGC AACTGCCCCG CGTCAGCGAG GAGATCGAAA CCCCTGTGGT CCGGCTCTCC
 73141 GACGTCATCG CGGAGCGCGG TATCGAGGCC ATCGGCCTGC TGAAGGTCGA CGTGGAGAAG
 73201 AGCGAACGGC AGGTCTTCGC CGGCCTCGAG GACACCGACT GGCCCGGTAT CCGCCAGGTC
 73261 GTCGCGGAGG TCCACGACAT CGACGGCGCG CTCGAGGAGG TCCTACGCT GCTCCGCGG
 20 73321 CATGGCTTCA CCGTGGTCCG CGAGCAGGAA CCGCTGTTCC CCGGCACGGG CATCCACCAG
 73381 GTCGCCGCGC GCGGGGTGGC CGGTGAGCG CCGTCGGGGC CGCGGCCGTC CGCACCAGCG
 73441 GCGCGGCTGC GGACGGCGGC TCAGCCGGCG TCGGACAGTT CCTTGGGCG TGTGTCAGCG
 73501 CCCTTACACC CCAGCTTGCG GAACACGTTG GTGAGGTGCT GTTCCACCGT GCTGGAGGTG
 73561 ACGAACAGCT GGCTGGCGAT CTCCTTGTGT GTGCGCCCGA CCGCGGCGTG CGACGCCACC
 73621 CGCCGCTCCG CCTCGGTGAG CGATGTGATC CGCTGCGCCG GCGTCACGTC CTGGGTGCGG
 25 73681 TCCGCGTCCG AGGACTCCCC ACCGAGCCGC CGGAGGAGCG GCACGGCTCC GCACTGGGTC
 73741 GCGAGGTGCC GTGCGCGGCG GAACAGTCCC CGCGCACGGC TGTGCCGCCG GAGCATGCCG
 73801 CACGCTTCGC CCATGTGCGC GAGGACGCGG GCCAGCTCGT ACTGGTCGCG GCACATGATG
 73861 AGCAGATCGG CGGCCTCGTC GAGCAGTTCG ATCCGCTTGG CCGGCGGACT GTAGGCCGCC
 73921 TGCACCCGCA GCGTCATAC CCGCGCCCGG GACCCCATCG GCCGGGACAG CTGCTCGGAG
 30 73981 ATGACCTCA GCGCTCGTC ACGGCCGCGG CCGAGCAGCA GAAGCGCTTC GCGGCGTCCG
 74041 ACCCGCCACA GGGCCAGGCC CGGCACGTCG ACGGACCAGC GTCGCATCCG CTCGCCGAG
 74101 TCCCGGAACG CGTTGTACGC CGCCCGGTAC CGCCCGGCCG CGAGATGGTG TTGCCACCG
 74161 GCGCAGACCA TGTGAGTCC GAAGAGGCTG TCGGAGGTCT CCTCCGCAA CGGCTCGGCG
 74221 AGCCACCGCT CCGCCCGGTC CAGGTGCGCC AGTCCGATCG CCGCGGCCAC GGTGCTGCTC
 35 74281 AGCGGCAATG CGGCGGCCAT CCCCCAGGAG GGCACGACCC GGGGGGCGAG CGCGGCCTCG
 74341 CCGATTGCA CGGCGGCGGT CAGGTGCGCC CGGCGCAGCG CGGCCTCGG CCGGAACCCC
 74401 GCGTGGACCG CCTCGTCGCG CGGGTCCGC ATGTTGTGCT CACCGGCCAG CATTCGAC
 74461 CAGGACTGGA CGGCATCGGT GTCCTCGGCG TAGAGCAGGG CCAGCAACGC CATCATGGTC
 74521 GTGGTCCGGT CCGTCGTGAC CCGGGAGTGC TGGAGCACGT ACTCGGCTTT GGCTCGGCC
 40 74581 GTTTCGGACC AGCCGCGCAG CGCGTTGCTC AGGGCCTTGT CCGCGACGGC GCGGTGCCCG
 74641 ACGGCTCCGG AAAACGAGGC GACCTCGTCC TCGGCCGGCG GATCGGCCGG ACGCGGCCGA
 74701 TCGGCCGCGC CGGGATAGAT CAGCGCGAGG GACAGGTCCG CGACGCGCAG GTGCGCCCGG
 74761 CCCTGCTCGC TCGGGGCGGC GGAGCGCTGG CGCGCCAGGA CCTCGCGGC CTCGCGCGG
 74821 CGCCCGTCCA TCGCCAGCCA GCAGGCGAGC GACACGGCGT GCTCGCTGGA GAGGACCGT
 45 74881 TCCCGCGACG CGGTGAGCAG CTCGGGCACA TGCCGGCCGG ATCTGGCGGG ATCGCAGAGC
 74941 CGCTCGATGG CGGCGGTGTC GACGCGCAGT GCGGCGTGA CCGCGGGGTC GTCGAGGGCC
 75001 CGGTAGGCGA ACTCCAGGTA GGTGACGGCC TCGTCGAGCT CGCCGCGCAG GTGGTGCTCG
 75061 CGCGCGGCGT CGGTGAACAG CCGGCGGACC TCGGCGCGCT GCACCCGGCC GGTACCCATC
 50 75121 TGGTGGCGGG CGAGCACCTT GCTGGCCACG CCGCGGTCCC GCAGCAGTTC CAGCGCCAGC
 75181 TCGTCAGGC CACGCCGCTC GCGGCGGGAG AGGTGCTCGA GTACGACGGA GCGGGCCGCG
 75241 GGGTGGCGGA ACCGCCCTTC CCGCAGCAGC CGCCCTCGA CCAGCTGTTT GTGGGCTGCG
 75301 TCGACCGCCT CGGTGTCGAG GCCGGTCATC CGCTGGACGA GGGTGAGTTC GACACTCTCG
 75361 CCGAGCACGG CGGAAGCTCG GGCGACGCTC AGCGCGGCGG GGCCGCAACG ATAGAGCGAC
 75421 CCGAGGTAGG CGAGCCGGTA CGCCCGCCCC GCGACCACTT CCAGGCACCC TGAGGTCCGT
 55 75481 GTCCGTGCCT CCCGGATGTC GTCGATCAGG CCGTGGCCGA GGAGCAGGTT GCCCGCGGTC
 75541 CCGCGGAACG CCTGGGCCAC CAGTGCCTCG TCGCGTCTCT GGCCGAGGTG CCGGCGCAGC
 75601 AGTTCGGTGG TCTGCGCTC GGTGAGCGGG CGCAGCGCGA TCTCTGTA GTGCGCAGA
 75661 CTCAGCAGTG CCGCCCGGAA TTGGGAGTGG GCGGGCGTCG GCGGAGCAGC CTCGGTCAGC
 75721 ACGATGGCGA CACGGGCCCG GCTGATGCGG CGCGCGAGGT GGAGCAGGCA GCGCAGCGAC
 60 75781 GCGCGTCCG CGTGGTGCAC GTCGTCGATG CCGATCAGTA CCGGCGGCTC CGCGGCGAGC

SUBSTITUTE SHEET (RULE 26)

75841 GTCAGCACCG TCGGGGTGAG TTCGGTCCCC AGGCGGTTGT CGACGTCGGC CGGCAGGTTT
 75901 TCGCACGATG CCGTCAGCCG GACCAGCTCC GGTGTCCGGG CGGCCAGCTC GGGCTGGTCG
 75961 AGGAGCTGGC CGAGCATGCC GTACGGCAGG GCCCGCTCCT CCATGGAGCA CACCGCGCGA
 5 76021 AGGGTGACGA AGCCGGCCTT GGCCGCGGCG GCGTCGAGGA GTTCGGTCTT GCCCGAGGCG
 76081 ATCGGCCCGG TGACGGCGGC GACGACGCCC CGCCCGCCCC CCGCTCGGGT GAGCGCCCGG
 76141 TGGAGGGAAC CGAACTCGTC ATCGCGGGCG ATCAGGTCTG GGGGAGATAA GCGCGCTATC
 76201 ACGAATGGAA CTACCTCGCG ACCGTCTGGG AAACCCATAG GCATCACATG GCTTGTGTGAT
 76261 CTGTACGGCT GTGATTGAGC CTGGCGGGAT GCTGTGCTAC AGATGGGAAG ATGTGATCTA
 10 76321 GGGCCGTGCC GTTCCCTCAG GAGCCGACCG CCCCCGGGCG CACCCGCCGT ACCCCTGGG
 76381 CCACCAGCTC GCGGACCCGC TCCTGGTGGT CGACGAGGTA GAAGTGCCCG CCGGGGAAGA
 76441 CCTCCACCGT GGTCGGCGCG GTCGTGTGCC CGGCCCAGGC GTGGGCCTGC TCCACCGTCG
 76501 TCTTCGGATC GTCGTACCG ATGCACACCG TGATCGGCGT CTCCAGCGGC GGC CGGGGCT
 76561 CCCACCGGTA CGTCTCCGC CGTAGTAGT CCGCCCGCAA CGGCGCCAGG ATCAGCGCGC
 76621 GCATTTCTGTC GTCCGCCATC ACATCGGCGC TCGTCCCGCC GAGGCCGATG ACCSCCGCCA
 15 76681 GCAGCTCGTC GTCGGACGCG AGGTGGTCCT GGTCGGCGCG CGGCTGCGAC GGCGCCGCC
 76741 GGCCCGAGAC GATCAGGTGC GCCACCGGGA GCCGTGGGC CAGCTCGAAC GCGAGTGTGCG
 76801 CGCCCATGCT GTGGCCGAAC AGCACCAGCG GACGGTCCAG CCCCCGCTTC AACGCCTCGG
 76861 CCACGAGGCC GGCGAGAACA CGCAGGTCGC GCACCGCCTC CTCGTGCGCG CGSTCCTGGC
 76921 GGCCGGGGTA CTGCACGGCG TACACGTCCG CCACCGGGGC GAGCGCACGG GCCAGCGGAA
 20 76981 GGTAGAACGT CGCCGATCCG CCGCGTGGG GCAGCAGCAC CACCCGTACC GGGGCCTCGG
 77041 GCGTGGGGAA GAACTGCCGC AGCCAGAGTT CCGAGCTCAC CGCACCCCTC CGGCCGCGAC
 77101 CTGGGGAGCC CGGAACCGGG TGATCTCGGC CAAGTGCTTC TCCCGCATCT CCGGTCGGT
 77161 CACGCCCCAT CCCTCCTCCG GCGCCAGACA GAGGACGCCG ACTTTGCCGT TGTGCACATT
 77221 GCGATGCACA TCGCGCACCG CCGACCCGAC GTCGTGAGC GGGTAGGTCA CCGACAGCGT
 25 77281 CGGGTGCACC ATCCCCTTGC AGATCAGGCG GTTCGCCTCC CACGCCTCAC GATAGTTCGC
 77341 GAAGTGGGTA CCGATGATCC GCTTCACGGA CATCCACAGG TACCGATTGT CAAAGGCGTG
 77401 CTCGTATCCC GAGGTTGACG CGCAGGTGAC GATCGTGCCA CCCCAGCTG TCACGTAGAC
 77461 ACTCGCGCCG AACGTCGCGC GCCCGGGTG CTGAACACG ATGTCGGGAT CGTCACCGCC
 30 77521 GGTACAGTCC CGGATC

Those of skill in the art will recognize that, due to the degenerate nature of the
 genetic code, a variety of DNA compounds differing in their nucleotide sequences can be
 used to encode a given amino acid sequence of the invention. The native DNA sequence
 encoding the FK-520 PKS of *Streptomyces hygroscopicus* is shown herein merely to
 35 illustrate a preferred embodiment of the invention, and the present invention includes DNA
 compounds of any sequence that encode the amino acid sequences of the polypeptides and
 proteins of the invention. In similar fashion, a polypeptide can typically tolerate one or more
 amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or
 significant loss of a desired activity. The present invention includes such polypeptides with
 40 alternate amino acid sequences, and the amino acid sequences shown merely illustrate
 preferred embodiments of the invention.

The recombinant nucleic acids, proteins, and peptides of the invention are many and
 diverse. To facilitate an understanding of the invention and the diverse compounds and
 methods provided thereby, the following general description of the FK-520 PKS genes and
 45 modules of the PKS proteins encoded thereby is provided. This general description is
 followed by a more detailed description of the various domains and modules of the FK-520

PKS contained in and encoded by the compounds of the invention. In this description, reference to a heterologous PKS refers to any PKS other than the FK-520 PKS. Unless otherwise indicated, reference to a PKS includes reference to a portion of a PKS. Moreover, reference to a domain, module, or PKS includes reference to the nucleic acids encoding the same and vice-versa, because the methods and reagents of the invention provide or enable one to prepare proteins and the nucleic acids that encode them.

The FK-520 PKS is composed of three proteins encoded by three genes designated *fkba*, *fkbb*, and *fkbc*. The *fkba* ORF encodes extender modules 7 - 10 of the PKS. The *fkbb* ORF encodes the loading module (the CoA ligase) and extender modules 1 - 4 of the PKS. The *fkbc* ORF encodes extender modules 5 - 6 of the PKS. The *fkbp* ORF encodes the NRPS that attaches the pipecolic acid and cyclizes the FK-520 polyketide.

The loading module of the FK-520 PKS includes a CoA ligase, an ER domain, and an ACP domain. The starter building block or unit for FK-520 is believed to be a dihydroxycyclohexene carboxylic acid, which is derived from shikimate. The recombinant DNA compounds of the invention that encode the loading module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of methods and in a variety of compounds. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for the loading module of the heterologous PKS is replaced by the coding sequence for the FK-520 loading module, provides a novel PKS coding sequence. Examples of heterologous PKS coding sequences include the rapamycin, FK-506, rifamycin, and avermectin PKS coding sequences. In another embodiment, a DNA compound comprising a sequence that encodes the FK-520 loading module is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the loading module coding sequence is utilized in conjunction with a heterologous coding sequence. In this embodiment, the invention provides, for example, either replacing the CoA ligase with a different CoA ligase, deleting the ER, or replacing the ER with a different ER. In addition, or alternatively, the ACP can be replaced by another ACP. In similar fashion, the corresponding domains in another loading or extender module can be replaced by one or more domains of the FK-520 PKS. The resulting heterologous loading module coding sequence can be utilized in conjunction

with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide.

The first extender module of the FK-520 PKS includes a KS domain, an AT domain specific for methylmalonyl CoA, a DH domain, a KR domain, and an ACP domain. The recombinant DNA compounds of the invention that encode the first extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 first extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the first extender module of the FK-520 PKS or the latter is merely added to coding sequences for modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the first extender module of the FK-520 PKS is inserted into a DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or only a portion of the first extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting either the DH or KR or both; replacing the DH or KR or both with another DH or KR; and/or inserting an ER. In replacing or inserting KR, DH, and ER domains, it is often beneficial to replace the existing KR, DH, and ER domains with the complete set of domains desired from another module. Thus, if one desires to insert an ER domain, one may simply replace the existing KR and DH domains with a KR, DH, and ER set of domains from a module containing such domains. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a gene for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous first extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous

PKS can be replaced by one or more domains of the first extender module of the FK-520 PKS.

5 In an illustrative embodiment of this aspect of the invention, the invention provides recombinant PKSs and recombinant DNA compounds and vectors that encode such PKSs in which the KS domain of the first extender module has been inactivated. Such constructs are especially useful when placed in translational reading frame with the remaining modules and domains of an FK-520 or FK-520 derivative PKS. The utility of these constructs is that host cells expressing, or cell free extracts containing, the PKS encoded thereby can be fed or supplied with N-acylcysteamine thioesters of novel precursor molecules to prepare FK-520 derivatives. See U.S. patent application Serial No. 60/117,384, filed 27 Jan. 1999, and PCT patent publication Nos. US97/02358 and US99/03986, each of which is incorporated herein by reference.

15 The second extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the second extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 second extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the second extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the second extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

25 In another embodiment, all or a portion of the second extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these

replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous second extender module coding sequence
5 can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the second extender module of the FK-520 PKS.

The third extender module of the FK-520 PKS includes a KS, an AT specific for
10 malonyl CoA, a KR, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the third extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 third extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS.
15 The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the third extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the third extender module of the FK-520 PKS is inserted into a DNA
20 compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, all or a portion of the third extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the
25 malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting the KR and/or the inactive DH; replacing the KR with another KR; and/or inserting an active DH or an active DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding
30 sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous third extender module coding sequence can be utilized in conjunction with a coding sequence from a PKS that synthesizes FK-520, an

FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the third extender module of the FK-520 PKS.

5 The fourth extender module of the FK-520 PKS includes a KS, an AT that binds ethylmalonyl CoA, an inactive DH, and an ACP. The recombinant DNA compounds of the invention that encode the fourth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fourth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS.
10 The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fourth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the fourth extender module of the FK-520 PKS is inserted into a
15 DNA compound that comprises the remainder of the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fourth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the ethylmalonyl CoA
20 specific AT with a malonyl CoA, methylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or deleting the inactive DH, inserting a KR, a KR and an active DH, or a KR, an active DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of
25 the FK-520 PKS, a PKS for a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fourth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fourth extender
30 module of the FK-520 PKS.

As illustrative examples, the present invention provides recombinant genes, vectors, and host cells that result from the conversion of the FK-506 PKS to an FK-520 PKS and vice-versa. In one embodiment, the invention provides a recombinant set of FK-506 PKS

genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-520 PKS. This recombinant PKS can be used to produce FK-520 in recombinant host cells. In another embodiment, the invention provides a recombinant set of FK-520 PKS genes but in which the coding sequences for the fourth extender module or at least those for the AT domain in the fourth extender module have been replaced by those for the AT domain of the fourth extender module of the FK-506 PKS. This recombinant PKS can be used to produce FK-506 in recombinant host cells.

Other examples of hybrid PKS enzymes of the invention include those in which the AT domain of module 4 has been replaced with a malonyl specific AT domain to provide a PKS that produces 21-desethyl-FK520 or with a methylmalonyl specific AT domain to provide a PKS that produces 21-desethyl-21-methyl-FK520. Another hybrid PKS of the invention is prepared by replacing the AT and inactive KR domain of FK-520 extender module 4 with a methylmalonyl specific AT and an active KR domain, such as, for example, from module 2 of the DEBS or oleandolide PKS enzymes, to produce 21-desethyl-21-methyl-22-desoxo-22-hydroxy-FK520. The compounds produced by these hybrid PKS enzymes are neurotrophins.

The fifth extender module of the FK-520 PKS includes a KS, an AT that binds methylmalonyl CoA, a DH, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the fifth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 fifth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the fifth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS. In another embodiment, a DNA compound comprising a sequence that encodes the fifth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the fifth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA

specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one or both of the DH and KR; replacing any one or both of the DH and KR with either a KR and/or DH; and/or inserting an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous fifth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the fifth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH domain of the fifth extender module have been deleted or mutated to render the DH non-functional. In one such mutated gene, the KR and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-19 to C-20 double bond of FK-520 and has a C-20 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant fifth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this fifth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (lacking the C-19 to C-20 double bond of FK-506 and having a C-20 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH domain of module 5 has been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The sixth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds

of the invention that encode the sixth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 sixth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the sixth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the sixth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the sixth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous sixth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the sixth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the DH and ER domains of the sixth extender module have been deleted or mutated to render them non-functional. In one such mutated gene, the KR, ER, and DH coding sequences are replaced with those encoding only a KR domain from another PKS gene. This can also be accomplished by simply replacing the coding sequences for extender module six with those for an extender module having a methylmalonyl specific AT and only a KR domain from a heterologous PKS gene, such as,

for example, the coding sequences for extender module two encoded by the *eryAI* gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that has a C-18 hydroxyl group. Such analogs are preferred neurotrophins, because they have little or no immunosuppressant activity. This recombinant sixth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this sixth extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (having a C-18 hydroxyl group) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the DH and ER domains of module 6 have been deleted or otherwise rendered inactive and thus produces this novel polyketide.

The seventh extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds of the invention that encode the seventh extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 seventh extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the seventh extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the seventh extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the seventh extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or

malonyl CoA specific AT; deleting the KR, the DH, and/or the ER; and/or replacing the KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous seventh extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the seventh extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the seventh extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-15 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant seventh extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this seventh extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-15-desmethoxy) FK-506 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 7 has been replaced and thus produces this novel polyketide.

In another illustrative embodiment, the present invention provides a hybrid PKS in which the AT and KR domains of module 7 of the FK-520 PKS are replaced by a methylmalonyl specific AT domain and an inactive KR domain, such as, for example, the AT and KR domains of extender module 6 of the rapamycin PKS. The resulting hybrid PKS produces 15-desmethoxy-15-methyl-16-oxo-FK-520, a neurotrophin compound.

The eighth extender module of the FK-520 PKS includes a KS, an AT specific for 2-hydroxymalonyl CoA, a KR, and an ACP. The recombinant DNA compounds of the invention that encode the eighth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 eighth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the eighth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the eighth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the eighth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the 2-hydroxymalonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or malonyl CoA specific AT; deleting or replacing the KR; and/or inserting a DH or a DH and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous eighth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the eighth extender module of the FK-520 PKS.

In an illustrative embodiment, the present invention provides a set of recombinant FK-520 PKS genes in which the coding sequences for the AT domain of the eighth extender module has been replaced with those encoding an AT domain for malonyl, methylmalonyl, or ethylmalonyl CoA from another PKS gene. The resulting PKS genes code for the expression of an FK-520 PKS that produces an FK-520 analog that lacks the C-13 methoxy group, having instead a hydrogen, methyl, or ethyl group at that position, respectively. Such

analogs are preferred, because they are more slowly metabolized than FK-520. This recombinant eighth extender module coding sequence can be combined with other coding sequences to make additional compounds of the invention. In an illustrative embodiment, the present invention provides a recombinant FK-520 PKS that contains both this eighth
5 extender module and the recombinant fourth extender module described above that comprises the coding sequence for the fourth extender module AT domain of the FK-506 PKS. The invention also provides recombinant host cells derived from FK-506 producing host cells that have been mutated to prevent production of FK-506 but that express this recombinant PKS and so synthesize the corresponding (C-13-desmethoxy) FK-506
10 derivative. In another embodiment, the present invention provides a recombinant FK-506 PKS in which the AT domain of module 8 has been replaced and thus produces this novel polyketide.

The ninth extender module of the FK-520 PKS includes a KS, an AT specific for methylmalonyl CoA, a KR, a DH, an ER, and an ACP. The recombinant DNA compounds
15 of the invention that encode the ninth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 ninth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of
20 the heterologous PKS is either replaced by that for the ninth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the ninth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the
25 FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion of the ninth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the methylmalonyl CoA specific AT with a malonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific
30 AT; deleting any one, two, or all three of the KR, DH, and ER; and/or replacing any one, two, or all three of the KR, DH, and ER with another KR, DH, and/or ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can

originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous ninth extender module coding sequence can be utilized in conjunction with a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a module of a heterologous PKS can be replaced by one or more domains of the ninth extender module of the FK-520 PKS.

The tenth extender module of the FK-520 PKS includes a KS, an AT specific for malonyl CoA, and an ACP. The recombinant DNA compounds of the invention that encode the tenth extender module of the FK-520 PKS and the corresponding polypeptides encoded thereby are useful for a variety of applications. In one embodiment, a DNA compound comprising a sequence that encodes the FK-520 tenth extender module is inserted into a DNA compound that comprises the coding sequence for a heterologous PKS. The resulting construct, in which the coding sequence for a module of the heterologous PKS is either replaced by that for the tenth extender module of the FK-520 PKS or the latter is merely added to coding sequences for the modules of the heterologous PKS, provides a novel PKS coding sequence. In another embodiment, a DNA compound comprising a sequence that encodes the tenth extender module of the FK-520 PKS is inserted into a DNA compound that comprises the coding sequence for the remainder of the FK-520 PKS or a recombinant FK-520 PKS that produces an FK-520 derivative.

In another embodiment, a portion or all of the tenth extender module coding sequence is utilized in conjunction with other PKS coding sequences to create a hybrid module. In this embodiment, the invention provides, for example, either replacing the malonyl CoA specific AT with a methylmalonyl CoA, ethylmalonyl CoA, or 2-hydroxymalonyl CoA specific AT; and/or inserting a KR, a KR and DH, or a KR, DH, and an ER. In addition, the KS and/or ACP can be replaced with another KS and/or ACP. In each of these replacements or insertions, the heterologous KS, AT, DH, KR, ER, or ACP coding sequence can originate from a coding sequence for another module of the FK-520 PKS, from a coding sequence for a PKS that produces a polyketide other than FK-520, or from chemical synthesis. The resulting heterologous tenth extender module coding sequence can be utilized in conjunction with a coding sequence for a PKS that synthesizes FK-520, an FK-520 derivative, or another polyketide. In similar fashion, the corresponding domains in a

module of a heterologous PKS can be replaced by one or more domains of the tenth extender module of the FK-520 PKS.

The FK-520 polyketide precursor produced by the action of the tenth extender module of the PKS is then attached to pipecolic acid and cyclized to form FK-520. The enzyme FkbP is the NRPS like enzyme that catalyzes these reactions. FkbP also includes a thioesterase activity that cleaves the nascent FK-520 polyketide from the NRPS. The present invention provides recombinant DNA compounds that encode the *fkbP* gene and so provides recombinant methods for expressing the *fkbP* gene product in recombinant host cells. The recombinant *fkbP* genes of the invention include those in which the coding sequence for the adenylation domain has been mutated or replaced with coding sequences from other NRPS like enzymes so that the resulting recombinant FkbP incorporates a moiety other than pipecolic acid. For the construction of host cells that do not naturally produce pipecolic acid, the present invention provides recombinant DNA compounds that express the enzymes that catalyze at least some of the biosynthesis of pipecolic acid (see Nielsen *et al.*, 1991, *Biochem.* 30: 5789-96). The *fkbL* gene encodes a homolog of RapL, a lysine cyclodeaminase responsible in part for producing the pipecolate unit added to the end of the polyketide chain. The *fkbB* and *fkbL* recombinant genes of the invention can be used in heterologous hosts to produce compounds such as FK-520 or, in conjunction with other PKS or NRPS genes, to produce known or novel polyketides and non-ribosomal peptides.

The present invention also provides recombinant DNA compounds that encode the P450 oxidase and methyltransferase genes involved in the biosynthesis of FK-520. Figure 2 shows the various sites on the FK-520 polyketide core structure at which these enzymes act. By providing these genes in recombinant form, the present invention provides recombinant host cells that can produce FK-520. This is accomplished by introducing the recombinant PKS, P450 oxidase, and methyltransferase genes into a heterologous host cell. In a preferred embodiment, the heterologous host cell is *Streptomyces coelicolor* CH999 or *Streptomyces lividans* K4-114, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference. In addition, by providing recombinant host cells that express only a subset of these genes, the present invention provides methods for making FK-520 precursor compounds not readily obtainable by other means.

In a related aspect, the present invention provides recombinant DNA compounds and vectors that are useful in generating, by homologous recombination, recombinant host

cells that produce FK-520 precursor compounds. In this aspect of the invention, a native host cell that produces FK-520 is transformed with a vector (such as an SCP2* derived vector for *Streptomyces* host cells) that encodes one or more disrupted genes (i.e., a hydroxylase, a methyltransferase, or both) or merely flanking regions from those genes.

5 When the vector integrates by homologous recombination, the native, functional gene is deleted or replaced by the non-functional recombinant gene, and the resulting host cell thus produces an FK-520 precursor. Such host cells can also be complemented by introduction of a modified form of the deleted or mutated non-functional gene to produce a novel compound.

10 In one important embodiment, the present invention provides a hybrid PKS and the corresponding recombinant DNA compounds that encode those hybrid PKS enzymes. For purposes of the present invention a hybrid PKS is a recombinant PKS that comprises all or part of one or more modules and thioesterase/cyclase domain of a first PKS and all or part of one or more modules, loading module, and thioesterase/cyclase domain of a second PKS.

15 In one preferred embodiment, the first PKS is all or part of the FK-520 PKS, and the second PKS is only a portion or all of a non-FK-520 PKS.

One example of the preferred embodiment is an FK-520 PKS in which the AT domain of module 8, which specifies a hydroxymalonyl CoA and from which the C-13 methoxy group of FK-520 is derived, is replaced by an AT domain that specifies a malonyl, methylmalonyl, or ethylmalonyl CoA. Examples of such replacement AT domains include the AT domains from modules 3, 12, and 13 of the rapamycin PKS and from modules 1 and 2 of the erythromycin PKS. Such replacements, conducted at the level of the gene for the PKS, are illustrated in the examples below. Another illustrative example of such a hybrid PKS includes an FK-520 PKS in which the natural loading module has been replaced with a loading module of another PKS. Another example of such a hybrid PKS is an FK-520 PKS in which the AT domain of module three is replaced with an AT domain that binds methylmalonyl CoA.

25 In another preferred embodiment, the first PKS is most but not all of a non-FK-520 PKS, and the second PKS is only a portion or all of the FK-520 PKS. An illustrative example of such a hybrid PKS includes an erythromycin PKS in which an AT specific for methylmalonyl CoA is replaced with an AT from the FK-520 PKS specific for malonyl CoA.

Those of skill in the art will recognize that all or part of either the first or second PKS in a hybrid PKS of the invention need not be isolated from a naturally occurring source. For example, only a small portion of an AT domain determines its specificity. See U.S. provisional patent application Serial No. 60/091,526, incorporated herein by reference.

5 The state of the art in DNA synthesis allows the artisan to construct *de novo* DNA compounds of size sufficient to construct a useful portion of a PKS module or domain. For purposes of the present invention, such synthetic DNA compounds are deemed to be a portion of a PKS.

Thus, the hybrid modules of the invention are incorporated into a PKS to provide a hybrid PKS of the invention. A hybrid PKS of the invention can result not only:

(i) from fusions of heterologous domain (where heterologous means the domains in that module are from at least two different naturally occurring modules) coding sequences to produce a hybrid module coding sequence contained in a PKS gene whose product is incorporated into a PKS,

15 but also:

(ii) from fusions of heterologous module (where heterologous module means two modules are adjacent to one another that are not adjacent to one another in naturally occurring PKS enzymes) coding sequences to produce a hybrid coding sequence contained in a PKS gene whose product is incorporated into a PKS,

20 (iii) from expression of one or more FK-520 PKS genes with one or more non-FK-520 PKS genes, including both naturally occurring and recombinant non-FK-520 PKS genes, and

(iv) from combinations of the foregoing.

Various hybrid PKSs of the invention illustrating these various alternatives are described herein.

25 Examples of the production of a hybrid PKS by co-expression of PKS genes from the FK-520 PKS and another non-FK-520 PKS include hybrid PKS enzymes produced by coexpression of FK-520 and rapamycin PKS genes. Preferably, such hybrid PKS enzymes are produced in recombinant *Streptomyces* host cells that produce FK-520 or FK-506 but have been mutated to inactivate the gene whose function is to be replaced by the rapamycin PKS gene introduced to produce the hybrid PKS. Particular examples include (i) replacement of the *fkfC* gene with the *rapB* gene; and (ii) replacement of the *fkfA* gene with the *rapC* gene. The latter hybrid PKS produces 13,15-didesmethoxy-FK-520, if the host cell

30

is an FK-520 producing host cell, and 13,15-didesmethoxy-FK-506, if the host cell is an FK-506 producing host cell. The compounds produced by these hybrid PKS enzymes are immunosuppressants and neurotrophins but can be readily modified to act only as neurotrophins, as described in Example 6, below.

5 Other illustrative hybrid PKS enzymes of the invention are prepared by replacing the *fkba* gene of an FK-520 or FK-506 producing host cell with a hybrid *fkba* gene in which:
(a) the extender module 8 through 10, inclusive, coding sequences have been replaced by the coding sequences for extender modules 12 to 14, inclusive, of the rapamycin PKS; and
(b) the module 8 coding sequences have been replaced by the module 8 coding sequence of
10 the rifamycin PKS. When expressed with the other, naturally occurring FK-520 or FK-506 PKS genes and the genes of the modification enzymes, the resulting hybrid PKS enzymes produce, respectively, (a) 13-desmethoxy-FK-520 or 13-desmethoxy-FK-506; and (b) 13-desmethoxy-13-methyl-FK-520 or 13-desmethoxy-13-methyl-FK-506. In a preferred embodiment, these recombinant PKS genes of the invention are introduced into the
15 producing host cell by a vector such as pHU204, which is a plasmid pRM5 derivative that has the well-characterized SCP2* replicon, the *colE1* replicon, the *tsr* and *bla* resistance genes, and a *cos* site. This vector can be used to introduce the recombinant *fkba* replacement gene in an FK-520 or FK-506 producing host cell (or a host cell derived therefrom in which the endogenous *fkba* gene has either been rendered inactive by
20 mutation, deletion or homologous recombination with the gene that replaces it) to produce the desired hybrid PKS.

In constructing hybrid PKSs of the invention, certain general methods may be helpful. For example, it is often beneficial to retain the framework of the module to be altered to make the hybrid PKS. Thus, if one desires to add DH and ER functionalities to a
25 module, it is often preferred to replace the KR domain of the original module with a KR, DH, and ER domain-containing segment from another module, instead of merely inserting DH and ER domains. One can alter the stereochemical specificity of a module by replacement of the KS domain with a KS domain from a module that specifies a different stereochemistry. See Lau *et al.*, 1999, "Dissecting the role of acyltransferase domains of
30 modular polyketide synthases in the choice and stereochemical fate of extender units," *Biochemistry* 38(5):1643-1651, incorporated herein by reference. Stereochemistry can also be changed by changing the KR domain. Also, one can alter the specificity of an AT domain by changing only a small segment of the domain. See Lau *et al.*, *supra*. One can

also take advantage of known linker regions in PKS proteins to link modules from two different PKSs to create a hybrid PKS. See Gokhale *et al.*, 16 Apr. 1999, "Dissecting and Exploiting Intermodular Communication in Polyketide Synthases," *Science* 284: 482-485, incorporated herein by reference.

5 The following Table lists references describing illustrative PKS genes and corresponding enzymes that can be utilized in the construction of the recombinant PKSs and the corresponding DNA compounds that encode them of the invention. Also presented are various references describing tailoring enzymes and corresponding genes that can be employed in accordance with the methods of the present invention.

10 **Avermectin**

U.S. Pat. No. 5,252,474 to Merck.

MacNeil *et al.*, 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256, A Comparison of the Genes Encoding the Polyketide Synthases for Avermectin, Erythromycin, and Nemadectin.

15 MacNeil *et al.*, 1992, *Gene* 115: 119-125, Complex Organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase.

Ikeda *et al.*, Aug. 1999, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, *Proc. Natl. Acad. Sci. USA* 96: 9509-9514.

20 **Candicidin (FR008)**

Hu *et al.*, 1994, *Mol. Microbiol.* 14: 163-172.

Epothilone

U.S. Pat. App. Serial No. 60/130,560, filed 22 April 1999.

Erythromycin

25 PCT Pub. No. 93/13663 to Abbott.

US Pat. No. 5,824,513 to Abbott.

Donadio *et al.*, 1991, *Science* 252:675-9.

Cortes *et al.*, 8 Nov. 1990, *Nature* 348:176-8, An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of

30 *Saccharopolyspora erythraea*.

Glycosylation Enzymes

PCT Pat. App. Pub. No. 97/23630 to Abbott.

FK-506

Motamedi *et al.*, 1998, The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506, *Eur. J. biochem.* 256: 528-534.

Motamedi *et al.*, 1997, Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506, *Eur. J. Biochem.* 244: 74-80.

Methyltransferase

US 5,264,355, issued 23 Nov. 1993, Methylating enzyme from *Streptomyces* MA6858. 31-O-desmethyl-FK-506 methyltransferase.

Motamedi *et al.*, 1996, Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK-506 and FK-520, *J. Bacteriol.* 178: 5243-5248.

Streptomyces hygroscopicus

U.S. patent application Serial No. 09/154,083, filed 16 Sep. 1998.

Lovastatin

U.S. Pat. No. 5,744,350 to Merck.

Narbomycin

U.S. patent application Serial No. 60/107,093, filed 5 Nov. 1998, and Serial No. 60/120,254, filed 16 Feb. 1999.

Nemadectin

MacNeil *et al.*, 1993, *supra*.

Niddamycin

Kakavas *et al.*, 1997, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, *J. Bacteriol.* 179: 7515-7522.

Oleandomycin

Swan *et al.*, 1994, Characterisation of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, *Mol. Gen. Genet.* 242: 358-362.

U.S. patent application Serial No. 60/120,254, filed 16 Feb. 1999.

Olano *et al.*, 1998, Analysis of a *Streptomyces antibioticus* chromosomal region involved in oleandomycin biosynthesis, which encodes two glycosyltransferases responsible for glycosylation of the macrolactone ring, *Mol. Gen. Genet.* 259(3): 299-308.

Picromycin

PCT patent application US99/15047, filed 2 Jul. 1999.

Xue *et al.*, 1998, Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the *pikC*-encoded cytochrome P450 in *Streptomyces venezuelae*, *Chemistry & Biology* 5(11): 661-667.

- 5 Xue *et al.*, Oct. 1998, A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity, *Proc. Natl. Acad. Sci. USA* 95: 12111 12116.

Platenolide

EP Pat. App. Pub. No. 791,656 to Lilly.

Rapamycin

- 10 Schwecke *et al.*, Aug. 1995, The biosynthetic gene cluster for the polyketide rapamycin, *Proc. Natl. Acad. Sci. USA* 92:7839-7843.

Aparicio *et al.*, 1996, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, *Gene* 169: 9-16.

15 Rifamycin

August *et al.*, 13 Feb. 1998, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S669, *Chemistry & Biology*, 5(2): 69-79.

Sorangium PKS

- 20 U.S. patent application Serial No. 09/144,085, filed 31 Aug. 1998.

Soraphen

U.S. Pat. No. 5,716,849 to Novartis.

- Schupp *et al.*, 1995, *J. Bacteriology* 177: 3673-3679. A *Sorangium cellulosum* (Myxobacterium) Gene Cluster for the Biosynthesis of the Macrolide Antibiotic Soraphen
25 A: Cloning, Characterization, and Homology to Polyketide Synthase Genes from Actinomycetes.

Spiramycin

U.S. Pat. No. 5,098,837 to Lilly.

Activator Gene

- 30 U.S. Pat. No. 5,514,544 to Lilly.

Tylosin

EP Pub. No. 791,655 to Lilly.

U.S. Pat. No. 5,876,991 to Lilly.

Kuhstoss *et al.*, 1996, *Gene* 183:231-6.. Production of a novel polyketide through the construction of a hybrid polyketide synthase.

Tailoring enzymes

Merson-Davies and Cundliffe, 1994, *Mol. Microbiol.* 13: 349-355. Analysis of five
5 tylosin biosynthetic genes from the *tylBA* region of the *Streptomyces fradiae* genome.

As the above Table illustrates, there are a wide variety of polyketide synthase genes that serve as readily available sources of DNA and sequence information for use in constructing the hybrid PKS-encoding DNA compounds of the invention. Methods for constructing hybrid PKS-encoding DNA compounds are described without reference to the
10 FK-520 PKS in PCT patent publication No. 98/51695; U.S. Patent Nos. 5,672,491 and 5,712,146 and U.S. patent application Serial Nos. 09/073,538, filed 6 May 1998, and 09/141,908, filed 28 Aug 1998, each of which is incorporated herein by reference.

The hybrid PKS-encoding DNA compounds of the invention can be and often are hybrids of more than two PKS genes. Moreover, there are often two or more modules in the
15 hybrid PKS in which all or part of the module is derived from a second (or third) PKS. Thus, as one illustrative example, the present invention provides a hybrid FK-520 PKS that contains the naturally occurring loading module and FkbP as well as modules one, two, four, six, seven, and eight, nine, and ten of the FK-520 PKS and further contains hybrid or heterologous modules three and five. Hybrid or heterologous module three contains an AT
20 domain that is specific of methylmalonyl CoA and can be derived for example, from the erythromycin or rapamycin PKS genes. Hybrid or heterologous module five contains an AT domain that is specific for malonyl CoA and can be derived for example, from the
picromycin or rapamycin PKS genes.

While an important embodiment of the present invention relates to hybrid PKS
25 enzymes and corresponding genes, the present invention also provides recombinant FK-520 PKS genes in which there is no second PKS gene sequence present but which differ from the FK-520 PKS gene by one or more deletions. The deletions can encompass one or more modules and/or can be limited to a partial deletion within one or more modules. When a deletion encompasses an entire module, the resulting FK-520 derivative is at least two
30 carbons shorter than the gene from which it was derived. When a deletion is within a module, the deletion typically encompasses a KR, DH, or ER domain, or both DH and ER domains, or both KR and DH domains, or all three KR, DH, and ER domains.

To construct a hybrid PKS or FK-520 derivative PKS gene of the invention, one can employ a technique, described in PCT Pub. No. 98/27203 and U.S. patent application Serial No. 08/989,332, filed 11 Dec. 1997, each of which is incorporated herein by reference, in which the large PKS gene is divided into two or more, typically three, segments, and each segment is placed on a separate expression vector. In this manner, each of the segments of the gene can be altered, and various altered segments can be combined in a single host cell to provide a recombinant PKS gene of the invention. This technique makes more efficient the construction of large libraries of recombinant PKS genes, vectors for expressing those genes, and host cells comprising those vectors.

Thus, in one important embodiment, the recombinant DNA compounds of the invention are expression vectors. As used herein, the term expression vector refers to any nucleic acid that can be introduced into a host cell or cell-free transcription and translation medium. An expression vector can be maintained stably or transiently in a cell, whether as part of the chromosomal or other DNA in the cell or in any cellular compartment, such as a replicating vector in the cytoplasm. An expression vector also comprises a gene that serves to produce RNA that is translated into a polypeptide in the cell or cell extract. Furthermore, expression vectors typically contain additional functional elements, such as resistance-conferring genes to act as selectable markers.

The various components of an expression vector can vary widely, depending on the intended use of the vector. In particular, the components depend on the host cell(s) in which the vector will be used or is intended to function. Vector components for expression and maintenance of vectors in *E. coli* are widely known and commercially available, as are vector components for other commonly used organisms, such as yeast cells and *Streptomyces* cells.

In a preferred embodiment, the expression vectors of the invention are used to construct recombinant *Streptomyces* host cells that express a recombinant PKS of the invention. Preferred *Streptomyces* host cell/vector combinations of the invention include *S. coelicolor* CH999 and *S. lividans* K4-114 host cells, which do not produce actinorhodin, and expression vectors derived from the pRM1 and pRM5 vectors, as described in U.S. Patent No. 5,830,750 and U.S. patent application Serial Nos. 08/828,898, filed 31 Mar. 1997, and 09/181,833, filed 28 Oct. 1998, each of which is incorporated herein by reference.

The present invention provides a wide variety of expression vectors for use in *Streptomyces*. For replicating vectors, the origin of replication can be, for example and without limitation, a low copy number vector, such as SCP2* (see Hopwood *et al.*, *Genetic Manipulation of Streptomyces: A Laboratory manual* (The John Innes Foundation, Norwich, U.K., 1985); Lydiate *et al.*, 1985, *Gene* 35: 223-235; and Kieser and Melton, 1988, *Gene* 65: 83-91, each of which is incorporated herein by reference), SLP1.2 (Thompson *et al.*, 1982, *Gene* 20: 51-62, incorporated herein by reference), and SG5(ts) (Muth *et al.*, 1989, *Mol. Gen. Genet.* 219: 341-348, and Bierman *et al.*, 1992, *Gene* 116: 43-49, each of which is incorporated herein by reference), or a high copy number vector, such as pIJ101 and pJV1 (see Katz *et al.*, 1983, *J. Gen. Microbiol.* 129: 2703-2714; Vara *et al.*, 1989, *J. Bacteriol.* 171: 5782-5781; and Servin-Gonzalez, 1993, *Plasmid* 30: 131-140, each of which is incorporated herein by reference). Generally, however, high copy number vectors are not preferred for expression of genes contained on large segments of DNA. For non-replicating and integrating vectors, it is useful to include at least an *E. coli* origin of replication, such as from pUC, p1P, p1I, and pBR. For phage based vectors, the phages phiC31 and KC515 can be employed (see Hopwood *et al.*, *supra*).

Typically, the expression vector will comprise one or more marker genes by which host cells containing the vector can be identified and/or selected. Useful antibiotic resistance conferring genes for use in *Streptomyces* host cells include the *ermE* (confers resistance to erythromycin and other macrolides and lincomycin), *tsr* (confers resistance to thiostrepton), *aadA* (confers resistance to spectinomycin and streptomycin), *aacC4* (confers resistance to apramycin, kanamycin, gentamicin, geneticin (G418), and neomycin), *hyg* (confers resistance to hygromycin), and *vph* (confers resistance to viomycin) resistance conferring genes.

The recombinant PKS gene on the vector will be under the control of a promoter, typically with an attendant ribosome binding site sequence. The present invention provides the endogenous promoters of the FK-520 PKS and related biosynthetic genes in recombinant form, and these promoters are preferred for use in the native hosts and in heterologous hosts in which the promoters function. A preferred promoter of the invention is the *fkfO* gene promoter, comprised in a sequence of about 270 bp between the start of the open reading frames of the *fkfO* and *fkfB* genes. The *fkfO* promoter is believed to be bi-directional in that it promotes transcription of the genes *fkfO*, *fkfP*, and *fkfA* in one direction and *fkfB*, *fkfC*, and *fkfL* in the other. Thus, in one aspect, the present invention

provides a recombinant expression vector comprising the promoter of the *fkbO* gene of an FK-520 producing organism positioned to transcribe a gene other than *fkbO*. In a preferred embodiment the transcribed gene is an FK-520 PKS gene. In another preferred embodiment the transcribed gene is a gene that encodes a protein comprised in a hybrid PKS.

5 Heterologous promoters can also be employed and are preferred for use in host cells in which the endogenous FK-520 PKS gene promoters do not function or function poorly. A preferred heterologous promoter is the *actI* promoter and its attendant activator gene *actII-ORF4*, which is provided in the pRM1 and pRM5 expression vectors, *supra*. This promoter is activated in the stationary phase of growth when secondary metabolites are normally
10 synthesized. Other useful *Streptomyces* promoters include without limitation those from the *ermE* gene and the *melC1* gene, which act constitutively, and the *tipA* gene and the *merA* gene, which can be induced at any growth stage. In addition, the T7 RNA polymerase system has been transferred to *Streptomyces* and can be employed in the vectors and host cells of the invention. In this system, the coding sequence for the T7 RNA polymerase is
15 inserted into a neutral site of the chromosome or in a vector under the control of the inducible *merA* promoter, and the gene of interest is placed under the control of the T7 promoter. As noted above, one or more activator genes can also be employed to enhance the activity of a promoter. Activator genes in addition to the *actII-ORF4* gene discussed above include *dnrI*, *redD*, and *ptpA* genes (see U.S. patent application Serial No. 09/181,833,
20 *supra*) to activate promoters under their control.

In addition to providing recombinant DNA compounds that encode the FK-520 PKS, the present invention also provides DNA compounds that encode the ethylmalonyl CoA and 2-hydroxymalonyl CoA utilized in the synthesis of FK-520. Thus, the present invention also provides recombinant host cells that express the genes required for the
25 biosynthesis of ethylmalonyl CoA and 2-hydroxymalonyl CoA. Figures 3 and 4 show the location of these genes on the cosmids of the invention and the biosynthetic pathway that produces ethylmalonyl CoA.

For 2-hydroxymalonyl CoA biosynthesis, the *fkbH*, *fkbI*, *fkbJ*, and *fkbK* genes are sufficient to confer this ability on *Streptomyces* host cells. For conversion of 2-
30 hydroxymalonyl to 2-methoxymalonyl, the *fkbG* gene is also employed. While the complete coding sequence for *fkbH* is provided on the cosmids of the invention, the sequence for this gene provided herein may be missing a T residue, based on a comparison made with a similar gene cloned from the ansamitocin gene cluster by Dr. H. Floss. Where the sequence

herein shows one T, there may be two, resulting in an extension of the *fkfH* reading frame to encode the amino acid sequence:

MTIVKCLVWDLNLTWRGTVLEDDEVVLTDEIREVITTLDDRGILQAVASKNDHD
 LAWERLERLGVAEYFVLARIGWGPKSQSVREIATELNFAPTTIAFIDDQPAERA EVA
 5 FHLPEVRCYPAEQAATLLSLPEFSPPVSTVDSRRRLMYQAGFARDQAREAYSGPD
 EDFLRSLDLSMTIAPAGEEELSRVEELTLRTSQMNATGVHYSDADLRALLTDP AHE
 VLVVTMGDRFGPHGAVGIILLEKKPSTWHLKLLATSCRVVVSFGAGATILNWLTDQG
 ARAG AHLVADFRRTDRNRMM EIAYRFAGFADSDCPCVSEVAGASAAGVERLHLEP
 SARPAPTTTLTAADIAPVTVSAAG.

10 For ethylmalonyl CoA biosynthesis, one requires only a crotonyl CoA reductase, which can be supplied by the host cell but can also be supplied by recombinant expression of the *fkfS* gene of the present invention. To increase yield of ethylmalonyl CoA, one can also express the *fkfE* and *fkfU* genes as well. While such production can be achieved using only the recombinant genes above, one can also achieve such production by placing into the
 15 recombinant host cell a large segment of the DNA provided by the cosmids of the invention. Thus, for 2-hydroxymalonyl and 2-methoxymalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the left side of the FK-520 PKS genes shown in Figure 1. For ethylmalonyl CoA biosynthesis, one can simply provide the cells with the segment of DNA located on the right side of the FK-520 PKS genes shown in
 20 Figure 1 or, alternatively, both the right and left segments of DNA.

The recombinant DNA expression vectors that encode these genes can be used to construct recombinant host cells that can make these important polyketide building blocks from cells that otherwise are unable to produce them. For example, *Streptomyces coelicolor* and *Streptomyces lividans* do not synthesize ethylmalonyl CoA or 2-hydroxymalonyl CoA.
 25 The invention provides methods and vectors for constructing recombinant *Streptomyces coelicolor* and *Streptomyces lividans* that are able to synthesize either or both ethylmalonyl CoA and 2-hydroxymalonyl CoA. These host cells are thus able to make polyketides, those requiring these substrates, that cannot otherwise be made in such cells.

In a preferred embodiment, the present invention provides recombinant
 30 *Streptomyces* host cells, such as *S. coelicolor* and *S. lividans*, that have been transformed with a recombinant vector of the invention that codes for the expression of the ethylmalonyl CoA biosynthetic genes. The resulting host cells produce ethylmalonyl CoA and so are preferred host cells for the production of polyketides produced by PKS enzymes that

comprise one or more AT domains specific for ethylmalonyl CoA. Illustrative PKS enzymes of this type include the FK-520 PKS and a recombinant PKS in which one or more AT domains is specific for ethylmalonyl CoA.

5 In a related embodiment, the present invention provides *Streptomyces* host cells in which one or more of the ethylmalonyl or 2-hydroxymalonyl biosynthetic genes have been deleted by homologous recombination or rendered inactive by mutation. For example, deletion or inactivation of the *fkfG* gene can prevent formation of the methoxyl groups at C-13 and C-15 of FK-520 (or, in the corresponding FK-506 producing cell, FK-506), leading to the production of 13,15-didesmethoxy-13,15-dihydroxy-FK-520 (or, in the
10 corresponding FK-506 producing cell, 13,15-didesmethoxy-13,15-dihydroxy-FK-506). If the *fkfG* gene product acts on 2-hydroxymalonyl and the resulting 2-methoxymalonyl substrate is required for incorporation by the PKS, the AT domains of modules 7 and 8 may bind malonyl CoA and methylmalonyl CoA. Such incorporation results in the production of a mixture of polyketides in which the methoxy groups at C-13 and C-15 of FK-520 (or FK-
15 506) are replaced by either hydrogen or methyl.

This possibility of non-specific binding results from the construction of a hybrid PKS of the invention in which the AT domain of module 8 of the FK-520 PKS replaced the AT domain of module 6 of DEBS. The resulting PKS produced, in *Streptomyces lividans*, 6-dEB and 2-desmethyl-6-dEB, indicating that the AT domain of module 8 of the FK-520
20 PKS could bind malonyl CoA and methylmalonyl CoA substrates. Thus, one could possibly also prepare the 13,15-didesmethoxy-FK-520 and corresponding FK-506 compounds of the invention by deleting or otherwise inactivating one or more or all of the genes required for 2-hydroxymalonyl CoA biosynthesis, i.e., the *fkfH*, *fkfI*, *fkfJ*, and *fkfK* genes. In any event, the deletion or inactivation of one or more biosynthetic genes required for
25 ethylmalonyl and/or 2-hydroxymalonyl production prevents the formation of polyketides requiring ethylmalonyl and/or 2-hydroxymalonyl for biosynthesis, and the resulting host cells are thus preferred for production of polyketides that do not require the same.

The host cells of the invention can be grown and fermented under conditions known in the art for other purposes to produce the compounds of the invention. See, e.g., U.S.
30 Patent Nos. 5,194,378; 5,116,756; and 5,494,820, incorporated herein by reference, for suitable fermentation processes. The compounds of the invention can be isolated from the fermentation broths of these cultured cells and purified by standard procedures. Preferred compounds of the invention include the following compounds: 13-desmethoxy-FK-506; 13-

desmethoxy-FK-520; 13,15-didesmethoxy-FK-506; 13,15-didesmethoxy-FK-520; 13-desmethoxy-18-hydroxy-FK-506; 13-desmethoxy-18-hydroxy-FK-520; 13,15-didesmethoxy-18-hydroxy-FK-506; and 13,15-didesmethoxy-18-hydroxy-FK-520. These compounds can be further modified as described for tacrolimus and FK-520 in U.S. Patent
5 Nos. 5,225,403; 5,189,042; 5,164,495; 5,068,323; 4,980,466; and 4,920,218, incorporated herein by reference.

Other compounds of the invention are shown in Figure 8, Parts A and B. In Figure 8, Part A, illustrative C-32-substituted compounds of the invention are shown in two columns under the heading R. The substituted compounds are preferred for topical administration
10 and are applied to the dermis for treatment of conditions such as psoriasis. In Figure 8, Part B, illustrative reaction schemes for making the compounds shown in Figure 8, Part A, are provided. In the upper scheme in Figure 8, Part B, the C-32 substitution is a tetrazole moiety, illustrative of the groups shown in the left column under R in Figure 8, Part A. In the lower scheme in Figure 8, Part B, the C-32 substitution is a disubstituted amino group,
15 where R₃ and R₄ can be any group similar to the illustrative groups shown attached to the amine in the right column under R in Figure 8, Part A. While Figure 8 shows the C-32-substituted compounds in which the C-15-methoxy is present, the invention includes these C-32-substituted compounds in which C-15 is ethyl, methyl, or hydrogen. Also, while C-21 is shown as substituted with ethyl or allyl, the compounds of the invention includes the C-
20 32-substituted compounds in which C-21 is substituted with hydrogen or methyl.

To make these C-32-substituted compounds, Figure 8, Part B, provides illustrative reaction schemes. Thus, a selective reaction of the starting compound (see Figure 8, Part B, for an illustrative starting compound) with trifluoromethanesulfonic anhydride in the presence of a base yields the C-32 O-triflate derivative, as shown in the upper scheme of
25 Figure 8, Part B. Displacement of the triflate with 1H-tetrazole or triazole derivatives provides the C-32 tetrazole or triazole derivative. As shown in the lower scheme of Figure 8, Part B, reacting the starting compound with p-nitrophenylchloroformate yields the corresponding carbonate, which, upon displacement with an amino compound, provides the corresponding carbamate derivative.

30 The compounds can be readily formulated to provide the pharmaceutical compositions of the invention. The pharmaceutical compositions of the invention can be used in the form of a pharmaceutical preparation, for example, in solid, semisolid, or liquid form. This preparation contains one or more of the compounds of the invention as an active

ingredient in admixture with an organic or inorganic carrier or excipient suitable for external, enteral, or parenteral application. The active ingredient may be compounded, for example, with the usual non-toxic, pharmaceutically acceptable carriers for tablets, pellets, capsules, suppositories, solutions, emulsions, suspensions, and any other form suitable for use. Suitable formulation processes and compositions for the compounds of the present invention are described with respect to tacrolimus in U.S. Patent Nos. 5,939,427; 5,922,729; 5,385,907; 5,338,684; and 5,260,301, incorporated herein by reference. Many of the compounds of the invention contain one or more chiral centers, and all of the stereoisomers are included within the scope of the invention, as pure compounds as well as mixtures of stereoisomers. Thus the compounds of the invention may be supplied as a mixture of stereoisomers in any proportion.

The carriers which can be used include water, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate, talc, corn starch, keratin, colloidal silica, potato starch, urea, and other carriers suitable for use in manufacturing preparations, in solid, semi-solid, or liquified form. In addition, auxiliary stabilizing, thickening, and coloring agents and perfumes may be used. For example, the compounds of the invention may be utilized with hydroxypropyl methylcellulose essentially as described in U.S. Patent No. 4,916,138, incorporated herein by reference, or with a surfactant essentially as described in EPO patent publication No. 428,169, incorporated herein by reference.

Oral dosage forms may be prepared essentially as described by Hondo *et al.*, 1987, *Transplantation Proceedings XIX*, Supp. 6: 17-22, incorporated herein by reference. Dosage forms for external application may be prepared essentially as described in EPO patent publication No. 423,714, incorporated herein by reference. The active compound is included in the pharmaceutical composition in an amount sufficient to produce the desired effect upon the disease process or condition.

For the treatment of conditions and diseases relating to immunosuppression or neuronal damage, a compound of the invention may be administered orally, topically, parenterally, by inhalation spray, or rectally in dosage unit formulations containing conventional non-toxic pharmaceutically acceptable carriers, adjuvant, and vehicles. The term parenteral, as used herein, includes subcutaneous injections, and intravenous, intramuscular, and intrasternal injection or infusion techniques.

Dosage levels of the compounds of the present invention are of the order from about 0.01 mg to about 50 mg per kilogram of body weight per day, preferably from about 0.1 mg

to about 10 mg per kilogram of body weight per day. The dosage levels are useful in the treatment of the above-indicated conditions (from about 0.7 mg to about 3.5 mg per patient per day, assuming a 70 kg patient). In addition, the compounds of the present invention may be administered on an intermittent basis, i.e., at semi-weekly, weekly, semi-monthly, or
5 monthly intervals.

The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. For example, a formulation intended for oral administration to humans may contain from 0.5 mg to 5 g of active agent compounded with an appropriate
10 and convenient amount of carrier material, which may vary from about 5 percent to about 95 percent of the total composition. Dosage unit forms will generally contain from about 0.5 mg to about 500 mg of active ingredient. For external administration, the compounds of the invention can be formulated within the range of, for example, 0.00001% to 60% by weight, preferably from 0.001% to 10% by weight, and most preferably from about 0.005% to 0.8%
15 by weight. The compounds and compositions of the invention are useful in treating disease conditions using doses and administration schedules as described for tacrolimus in U.S. Patent Nos. 5,542,436; 5,365,948; 5,348,966; and 5,196,437, incorporated herein by reference. The compounds of the invention can be used as single therapeutic agents or in combination with other therapeutic agents. Drugs that can be usefully combined with
20 compounds of the invention include one or more immunosuppressant agents such as rapamycin, cyclosporin A, FK-506, or one or more neurotrophic agents.

It will be understood, however, that the specific dosage level for any particular patient will depend on a variety of factors. These factors include the activity of the specific compound employed; the age, body weight, general health, sex, and diet of the subject; the
25 time and route of administration and the rate of excretion of the drug; whether a drug combination is employed in the treatment; and the severity of the particular disease or condition for which therapy is sought.

A detailed description of the invention having been provided above, the following examples are given for the purpose of illustrating the present invention and shall not be
30 construed as being a limitation on the scope of the invention or claims.

Example 1

Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-520

The C-13 methoxyl group is introduced into FK-520 via an AT domain in extender module 8 of the PKS that is specific for hydroxymalonyl and by methylation of the hydroxyl group by an S-adenosyl methionine (SAM) dependent methyltransferase.

Metabolism of FK-506 and FK-520 primarily involves oxidation at the C-13 position into an inactive derivative that is further degraded by host P450 and other enzymes. The present invention provides compounds related in structure to FK-506 and FK-520 that do not contain the C-13 methoxy group and exhibit greater stability and a longer half-life *in vivo*. These compounds are useful medicaments due to their immunosuppressive and neurotrophic activities, and the invention provides the compounds in purified form and as pharmaceutical compositions.

The present invention also provides the novel PKS enzymes that produce these novel compounds as well as the expression vectors and host cells that produce the novel PKS enzymes. The novel PKS enzymes include, among others, those that contain an AT domain specific for either malonyl CoA or methylmalonyl CoA in module 8 of the FK-506 and FK-520 PKS. This example describes the construction of recombinant DNA compounds that encode the novel FK-520 PKS enzymes and the transformation of host cells with those recombinant DNA compounds to produce the novel PKS enzymes and the polyketides produced thereby.

To construct an expression cassette for performing module 8 AT domain replacements in the FK-520 PKS, a 4.6 kb *Sph*I fragment from the FK-520 gene cluster was cloned into plasmid pLitmus 38 (a cloning vector available from New England Biolabs). The 4.6 kb *Sph*I fragment, which encodes the ACP domain of module 7 followed by module 8 through the KR domain, was isolated from an agarose gel after digesting the cosmid pKOS65-C31 with *Sph* I. The clone having the insert oriented so the single *Sac*I site was nearest to the *Spe*I end of the polylinker was identified and designated as plasmid pKOS60-21-67. To generate appropriate cloning sites, two linkers were ligated sequentially as follows. First, a linker was ligated between the *Spe*I and *Sac*I sites to introduce a *Bgl*II site at the 5' end of the cassette, to eliminate interfering polylinker sites, and to reduce the total insert size to 4.5 kb (the limit of the phage KC515). The ligation reactions contained 5 picomolar unphosphorylated linker DNA and 0.1 picomolar vector DNA, i.e., a 50-fold molar excess of linker to vector. The linker had the following sequence:

5'-CTAGTGGGCAGATCTGGCAGCT-3'
3'-ACCCGTCTAGACCG-5'

The resulting plasmid was designated pKOS60-27-1.

Next, a linker of the following sequence was ligated between the unique *Sph*I and *Afl*III sites of plasmid pKOS60-27-1 to introduce an *Nsi*I site at the 3' end of the module 8 cassette. The linker employed was:

5' -GGGATGCATGGC-3'
 3' -GTACCCCTACGTACCGAATT-5'

The resulting plasmid was designated pKOS60-29-55.

To allow in-frame insertions of alternative AT domains, sites were engineered at the 5' end (*Avr* II or *Nhe* I) and 3' end (*Xho* I) of the AT domain using the polymerase chain reaction (PCR) as follows. Plasmid pKOS60-29-55 was used as a template for the PCR and sequence 5' to the AT domain was amplified with the primers *Spe*Bgl-fwd and either *Avr*-rev or *Nhe*-rev:

*Spe*Bgl-fwd 5' -CGACTCACTAGTGGGCAGATCTGG-3'

Avr-rev 5' -CACGCCTAGGCCGGTCGGTCTCGGGCCAC-3'

Nhe-rev 5' -GCGGCTAGCTGCTCGCCCATCGCGGGATGC-3'

The PCR included, in a 50 µl reaction, 5 µl of 10x *Pfu* polymerase buffer (Stratagene), 5 µl 10x z-dNTP mixture (2 mM dATP, 2 mM dCTP, 2 mM dTTP, 1 mM dGTP, 1 mM 7-deaza-GTP), 5 µl DMSO, 2 µl of each primer (10 µM), 1 µl of template DNA (0.1 µg/µl), and 1 µl of cloned *Pfu* polymerase (Stratagene). The PCR conditions were 95°C for 2 min., 25 cycles at 95°C for 30 sec., 60°C for 30 sec., and 72°C for 4 min., followed by 4 min. at 72°C and a hold at 0°C. The amplified DNA products and the Litmus vectors were cut with the appropriate restriction enzymes (*Bgl*II and *Avr*II or *Spe*I and *Nhe*I), and cloned into either pLitmus 28 or pLitmus38 (New England Biolabs), respectively, to generate the constructs designated pKOS60-37-4 and pKOS60-37-2, respectively.

Plasmid pKOS60-29-55 was again used as a template for PCR to amplify sequence 3' to the AT domain using the primers *Bsr*Xho-fwd and *Nsi*Afl-rev:

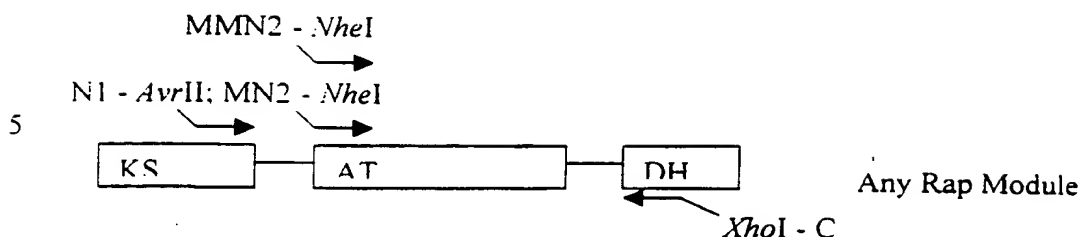
*Bsr*Xho-fwd 5' -GATGTACAGCTCGAGTCGGCACGCCCCGGCCGCATC-3'

*Nsi*Afl-rev 5' -CGACTCACTTAAGCCATGCATCC-3'

PCR conditions were as described above. The PCR fragment was cut with *Bsr*GI and *Afl*III, gel isolated, and ligated into pKOS60-37-4 cut with *Asp*718 and *Afl*III and inserted into pKOS60-37-2 cut with *Bsr*GI and *Afl*III, to give the plasmids pKOS60-39-1 and pKOS60-39-13, respectively. These two plasmids can be digested with *Avr*II and *Xho*I or *Nhe*I and *Xho*I, respectively, to insert heterologous AT domains specific for malonyl, methylmalonyl, ethylmalonyl, or other extender units.

Malonyl and methylmalonyl-specific AT domains were cloned from the rapamycin cluster using PCR amplification with a pair of primers that introduce an *AvrII* or *NheI* site at the 5' end and an *XhoI* site at the 3' end. The PCR conditions were as given above and the primer sequences were as follows:

- 5
RATN1 5'-ATCCTAGGCGGGCRGGYGTGTCGTCCTTCGG-3'
(3' end of Rap KS sequence and universal for malonyl and methylmalonyl CoA),
RATMN2 5'-ATGCTAGCCGCCGCGTTCCCCGTCTTCGCGCG-3'
(Rap AT shorter version 5'- sequence and specific for malonyl CoA),
10 RATMMN2 5'-ATGCTAGCGGATTCGTCGGTGGTGTTCGCCGA-3'
(Rap AT shorter version 5'- sequence and specific for methylmalonyl CoA), and
RATC 5'-ATCTCGAGCCAGTASCGCTGGTGYTGGAAGG-3'
(Rap DH 5'- sequence and universal for malonyl and methylmalonyl CoA).



Because of the high sequence similarity in each module of the rapamycin cluster, each primer was expected to prime any of the AT domains. PCR products representing ATs specific for malonyl or methylmalonyl extenders were identified by sequencing individual cloned PCR products. Sequencing also confirmed that the chosen clones contained no cloning artifacts. Examples of hybrid modules with the rapamycin AT12 and AT13 domains are shown in a separate figure.

The *AvrII*-*XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 12 of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below. The AT of rap module 12 is specific for incorporation of malonyl units.

```

20 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
   I W Q L A E A L L T L V R E S T
   GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCGACGGCGGC 100
   A A V L A G H V G G E D I P A T A A
   GTTCAAGGACCTCGGCATCGACTCGCTCACCGCGGTCCAGCTGCGCAACG 150
   F K D L G I D S L T A V Q L R N
   CCCTCACCGAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
   A L T E A T G V R L N A T A V F D
   TTCCCGACCCCGCACGTGCTCGCCGGAAGCTCGGCGACGAACTGACCGG 250
   F P T P H V L A G K L G D E L T G
   CACCCGCGCGCCCCGTGTCGCCCGGACCGCGGCCACGGCCGGTGCGCACG 300
   T R A P V V P R T A A T A G A H
   ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350
   D E P L A I V G M A C R L P G G V
   GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
   A S P E E L W H L V A S G T D A I
   CACGGAGTTCCCGACGGACCGCGGTGGGACGTCGACGCGATCTACGACC 450
   T E F P T D R G W D V D A I Y D
   CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
   P D P D A I G K T F V R H G G F L
   ACCGGCGCGACAGGCTTCGACGCGGCGTTCCTTCGGCATCAGCCCGCGCGA 550
   T G A T G F D A A F F G I S P R E
   GGCCCTCGCGATGGACCCGACGAGCGGGTGCTCCTGGAGACGTGCTGGG 600
   A L A M D P Q Q R V L L E T S W
   AGGCGTTTCGAAAGCGCCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650
   E A F E S A G I T P D S T R G S D
   ACCGGCGTGTTCGTGCGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700
   T G V F Y G A F S Y G Y G T G A D
   CACCGACGGCTTCGGGCGGACCGGCTCGCAGACCAAGTGTGCTCTCGGGCC 750
   T D G F G A T G S Q T S V L S G
50 GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTTCACGGTCGACACG 800

```

R L S Y F Y G L E G P A V T V D T
CGGTGTTCTGCTGCTGCTGGTGGCGCTGCACAGGCGGGGCGAGTCTGCTGCG 850
A C S S S L V A L H Q A G Q S L R
CTGCGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
S G E C S L A L V G G V T V M A
CTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 950
S P G G F V E F S R Q R G L A P D
GGCGGGGCGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000
G R A K A F G A G A D G T S F A E
GGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
G A G V L I V E R L S D A E R N
GTCACACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
G H T V L A V V R G S A V N Q D G
GCCTCCAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
A S N G L S A P N G P S Q E R V I
CCGGCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
R Q A L A N A G L T P A D V D A
TCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250
V E A H G T G T R L G D P I E A Q
GCGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1300
A V L A T Y G Q E R A T P L L L G
CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350
S L K S N I G H A Q A A S G V A
GCATCATCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400
G I I K M V Q A L R H G E L P P T
CTGCACGCGGACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
L H A D E P S P H V D W T A G A V
CGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
E L L T S A R P W P E T D R P R
GGGCGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1550
R A G V S S F G I S G T N A H V I
CTGGAAAGCGCACCCCCCACTCAGCCTGCGGACAACGCGGTGATCGAGCG 1600
L E S A P P T Q P A D N A V I E R
GGCAGCGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1650
A P E W V P L V I S A R T Q S A
TGACTGAGCAGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
L T E H E G R L R A Y L A A S P G
GTGGATATGCGGGCTGTGGCATCGACGCTGGCGATGACACGGTGGGTGTT 1750
V D M R A V A S T L A M T R S V F
CGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
E H P A V L L G D D T V T G T A
TGTCTGACCTCGGGCGGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1850
V S D P R A V F V F P G Q G S Q R
GCTGGCATGGGTGAGGAAGTGGCGCGCGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 1900
A G M G E E L A A A F P V F A R I
CCATCAGCAGGTGTGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1950
H Q Q W D L D V P D L E V N
AGACCGGTTACGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000
E T G Y A Q P A L F A M Q V A L F
GGGCTGCTGGAATCGTGGGGGTGTACGACCGGACGCGGTGATCGGCCATTG 2050
G L L E S W G V R P D A V I G H S
GCTGGGTGAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
V G E L A A A Y V S G V W S L E
ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2150
D A C T L V S A R A R L M Q A L P
CGGGGTGGGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2200
A G G V X V A V P V S E D E A R A
GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2250
V L G E G V E I A A V N G P S S
TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300

V V L S G D E A A V L Q A A E S L
GGGAAGTGGACGCGGCTGGCGACCAGCCACGCGTTCCATTCCGCCCCSTAT 2350
G K W T R L A T S H A F H S A F M
5 GGAACCCATGCTGGAGGASTTCCGGGCGGTCCCGGAAGGCGTGACCTACC 2400
E P M L E E F P A V A E G L T Y
GGACGCCCGCAGGTCTCCATGGCCGTTGGTGATCAGGTGACCACCGCTGAG 2450
E T P Q V S M A V G D Q V T T A E
TACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGC 2500
Y W V R Q V R D T V R F G E Q V A
10 CTCGTACGAGGACGCCGTGTTTCGTGCGAGCTGGGTGCCGACCGGTCACTGG 2550
S Y E D A V F V E L G A D R S L
CCCGCCTGGTGCAGCGGTGTGCGGATGCTGCACGGCGACCACGAAATCCAG 2600
A R L V D G V A M L H G D H E I Q
GCCGCGATCGGCGCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCGA 2650
15 A A I G A L A H L Y V N G V T V D
CTGGCCCCGCGCTCCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTC 2700
W P A L L G D A P A T R V L D L
CGACATACGCCTTCCAGCACCAGCGCTACTGGCTCGAGTCGGCAGCGCCG 2750
P T Y A F Q H Q R Y W L E S A R P
20 GCCGCATCCGACGCGGGGCCACCCCGTGCTGGGCTCCGGTATCGCCCTCGC 2800
A A S D A G H P V L G S G I A L A
CGGGTCGCGGGGCGGGTGTTCACGGGTTCGCTGCCGACCGGTGCCGACC 2850
G S P G R V F T G S V P T G A D
GCGCGGTGTTTCGTGCGCGAGCTGGCGCTGGCCGCCGCGGACGCGGTGAC 2900
25 R A V F V A E L A L A A D A V D
TGCGCCACGGTCGAGCGGCTCGACATCGCCTCCGTGCCCGGCCGCGCGGG 2950
C A T V E R L D I A S V P G R P G
CCATGGCCGGACGACCGTACAGACCTGGGTGCGACGAGCCGGCGGACGACG 3000
H G R T T V Q T W V D E P A D D
30 GCCGGCGCGCGTTTACCGTGCACACCCGCACCGGCGACGCCCCGTGGACG 3050
G R R F T V H T R T G D A P W T
CTGCACGCCGAGGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATGC 3100
L H A E G V L R P H G T A L P D A
GGCCGACGCCGAGTGGCCCCCACCAGGGCGCGGTGCCCGCGGACGGGCTGC 3150
35 A D A E W P P P G A V P A D G L
CGGGTGTGTGGCGCCGGGGGACCAGGTCTTCGCCGAGGCCGAGGTGGAC 3200
P G V W R R G D Q V F A E A E V D
GGACCGGACGGTTTTCGTGGTGACCCCGACCTGCTCGACGCGGTCTTCTC 3250
G P D G F V V H P D L L D A V F S
40 CGCGGTGCGCGACGGAAGCCGCCAGCCGGCCGGATGSCGCGACCTGACGG 3300
A V G D G S R Q P A G W R D L T
TGCACGCGTCGGACGCCACCGTACTGCGCGCCTGCCTCACCCGGCGCACCC 3350
V H A S D A T V L R A C L T R R T
GACGGAGCCATGGGATTCCCGCCCTTCGACGGCGCCGGCCTGCCGGTACT 3400
45 D G A M G F A A F D G A G L P V L
CACCGCGGAGGCGGTGACGCTGCGGGAGGTGGCGTCAACGTCCGGCTCCG 3450
T A E A V T L R E V A S P S G S
AGGAGTCGGACGGCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCG 3500
E E S D G L H R L E W L A V A E A
50 GTCTACGACGGTGACCTGCCGAGGGACATGTCCTGATCACCGCCGCCCA 3550
V Y D G D L P E G H V L I T A A H
CCCCGACGACCCGAGGACATACCCACCGCGCCACACCCGCGCCACCC 3600
P D D P E D I P T R A H T R A T
GCGTCTTGACCGCCCTGCAACACCACCTCACCACCACCGACCACACCCCTC 3650
55 R V L T A L Q H H L T T T D H T L
ATCGTCCACACCACCCACCGACCCCGCGGCGCCACCGTCACCGGCCTCAC 3700
I V H T T T D P A G A T V T G L T
CCCGACCGCCCGAGAACGACACCCCGACCGCATCCGCTCATCGAAACCG 3750
R T A Q N E H P H R I R L I E T
60 ACCACCCCGACACCCCGCTCCCGCTGGCCCAACTCGCCACCCCTCGACCAC 3800

SUBSTITUTE SHEET (RULE 26)

D H P H T P L P L A Q L A T L D H
 CCCCACCTCCGCTCACCACCCACCCCTCCACCACCCCCACCTCACCCTC 3850
 P H L R L T H H T L H H P H L T P
 CCTCCACACCCACCCACCCACCCACCCACCCCTCACCACCCGAACACG 3900
 5 L H T T T P P T T T P L N P E H
 CCATCATCATCACCAGGCGGCTCCGGCACCCCTCGCCGGCATCCTCGCCCGC 3950
 A I I I T G G S G T L A G I L A R
 CACCTGAACACCCACACCTACCTCCTCTCCCGCACCCACCCCGCA 4000
 H L N H P H T Y L L S R T P P P D
 10 CCCCACCCCGGACCCACCTCCCTGCGACGTCCGGCGACCCACCAAC 4050
 A T P G T H L P C D V G D P H Q
 TCGCCACCCCTCACCACATCCCCAACCCCTCACCGCCATCTTCCAC 4100
 L A T T L T H I P Q P L T A I F H
 ACCGCCGCCACCCCTCGACGACGGCATCCTCCACGCCCTCACCCTCGACCG 4150
 15 T A A T L D D G I L H A L T P D R
 CCTCACCACCGTCTCCACCCCAAAGCCAACGCCGCTGGCACCTGCACC 4200
 L T T V L H P K A N A A W H L H
 ACCTCACCCAAACCAACCCCTCACCACCTTCTGCTCTACTCCAGCGCC 4250
 H L T Q N Q P L T H F V L Y S S A
 20 GCCGCCGTCTCGGACGCCCGGACAAGGAACTACGCCGCCGCAACGC 4300
 A A V L G S P G Q G N Y A A A N A
 CTTCTCGACGCCCTCGCCACCCACCGCCACACCTCGGCCAACCCGCCA 4350
 F L D A L A T H R H T L G Q P A
 CCTCCATCGCCTGGGGCATGTGGCACACCACCAGCACCTCACCAGGACAA 4400
 25 T S I A W G M W H T T S T L T G Q
 CTGACGACGCCGACCGGGACCGCATCCGCCGCGGGGTTTCTCCCGAT 4450
 L D D A D R D R I R R G G F L P I
 CACGGACGACGAGGGCATGGGGATGCAT
 30 T D D E G

The *AvrII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

35 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
 Q L A E A L L T L V R E S T
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCCGCCACGGCGGC 100
 A A V L G H V G G E D I P A T A A
 GTTCAAGCCTCGGCATCGACTCGCTCACCGCCGTCCAGCTGCGCAACG 150
 40 F K D L G I D S L T A V Q L R N
 CCCTCACCAGGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
 A L T E A T G V R L N A T A V F D
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAACTGACCGG 250
 F P T F H V L A G K L G D E L T G
 45 CACCCGCCGCCCGTCTGTCGCCCGGACCGCGGCCACGGCCGGTGCGCACC 300
 T R A P V V P R T A A T A G A H
 ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGCGGGGTC 350
 D E P L A I V G M A C R L P G G V
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
 50 A S P E E L W H L V A S G T D A I
 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTGACGCGATCTACGACC 450
 T E F P T D R G W D V D A I Y D
 CGGACCCCGACGCGATCGGCAAGACCTTCTGTCGGGCACGGTGGCTTCCTC 500
 P C P D A I G K T F V R H G G F L
 55 ACCGGCGCGACAGGCTTCGACGCGCGTTCCTTGGCATCAGCCCGCGCGA 550
 T G A T G F D A A F F G I S P R E
 GGCCCTCGCGATGGACCCCGACGAGCGGGTGCTCCTGGAGACGTCTGGG 600

A L A M D P Q Q R V L L E T S W
AGGCGTTTCGAAAGGCGCGGCATCACCCCGGACTCGAACCCGCGGCAGCGAC 650
E A F E S A G I T P D S T R G S D
ACCGGGGTGTTCTGCGCGCGCTTCTCTACGGTTACCGCACCGGTGCGGA 700
5 T G V F Y S A F S Y G Y G T G A D
CACCGACGGCTTCCGCGCGACCGGCTCGCAGACCACTGTGCTCTCCGGCC 750
T D G F E A T G S Q T S V L S G
GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTCCACACG 800
R L S Y F Y G L E G P A V T V D T
10 CCGTGTCTGCTCTGCTGGTGGCGCTGCACCAGGCGCGGCAGTCTGCTGCG 850
A C S S S L V A L H Q A G Q S L R
CTCCGGCGAATGCTCGCTCGCCCTGGTCTGGCGGGCTCACGGTGATGGCGT 900
S G E C S L A L V G G V T V M A
CTCCCGCGCGGCTTCTGGAGTTCTCCCGGCAGCGCGGCTCTGCGCCGGAC 950
15 S P G G F V E F S R Q R G L A P D
GGCCGGGCGAAGGCGTTCCGCGCGGGTGGCGACGGCACGAGCTTCGCCGA 1000
G R A K A F G A G A D G T S F A E
GGGTGCCGGTGTCTGATCTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
G A V L V Q A L R L S D A E R N
20 GTCACACCGTCTGCGGGTCTGCTCGTGGTTCGGCGGTCAACCAGGATGGT 1100
G H T V L A V V R G S A V N Q D G
GCCTCCAACGGGCTGTGCGCGCGCAACGGGCGGTCTCAGGAGCGGGTGAT 1150
A S N G L S A P N G P S Q E R V I
CCGGCAGGCCCTGCGCAACGCGGGGCTCACCCCGGCGGACGTGGAACGCG 1200
25 R Q A L A N A G L T P A D V D A
TCGAGGCCCCACGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250
V E A H G T G T R L G D F I E A Q
GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTCTGGG 1300
A V L A T Y G Q E R A T P L L L G
30 CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCTGCCG 1350
S L K S N I G H A Q A A S G V A
GCATCAKCAAGATGGTGCAGGCCCTCCGGCACGGGAGCTGCCGCCGACG 1400
G I I K M V Q A L R H G E L P P T
CTGCACGCCGACGAGCCGTCGCGCACGTCGACTGGACGGCCGGCGCGCGT 1450
35 L H A D E P S P H V D W T A G A V
CGAACTGCTGACGTGCGCCCGGCCGTGGCCCCGAGACCGACCGCCTAGGC 1500
E L L T S A R P W P E T D R P R
GGGCGGGCGTGTCTCTCTCGGAGTCAGCGGCACCAACGCCACGTCATC 1550
R A G V S S F G V S G T N A H V I
40 CTGGAGAGCGCACCCGCCGCTCAGCCCGCGGAGGAGCGCAGCCTGTTGA 1600
L E S A P P A Q P A E E A Q P V E
GACGCCGGTGGTGGCCCTCGGATGTGCTGCGGCTGGTGATATCGGCCAAGA 1650
T P V V A S D V L P L V I S A K
CCGAGCCCGCCCTGACCGAACACGAAGACCGGCTGCGCGCCTACCTGGCG 1700
45 T Q P A L T E H E D R L R A Y L A
GCGTCCGCCCGGGCGGATATACGGGCTGTGGCATCGACGCTGGCGGTGAC 1750
A S P G A D I R A V A S T L A V T
ACGGTCCGGTGTTCGAGCACCGCGCCGTA CTCTTGAGATGACACCGTCA 1800
R S V F E H R A V L L G D D T V
50 CCGGCACCGCGGTGACCGACCCAGGATCGTGTTTGTCTTTCCCGGGCAG 1850
T G T A V T D P R I V F V F P G Q
GGGTGGCAGTGGCTGGGATGGGAGTGCAGTGCAGGATTCTGTCGGTGGT 1900
G W G L G M G S A L R D S S V V
GTTCGCGGAGCGGATGGCCGAGTGTGCGGCGGCGTTGCGCGAGTTCGTGG 1950
55 F A E R M A E C A A A L R E F V
ACTGGGATCTGTTACGGTCTGAGATGATCCGGCGGTGGTGGACCGGGTT 2000
D W E L F T V L D D P A V V D R V
GATGTGGTCCAGCCCGCTCTCTGGGCGATGATGGTTTCCCTGGCGCGGT 2050
D V V Q P A S W A M M V S L A A V
60 GTGGCAGGCGCGCCGTCTGCGGCGGATGCGGTGATCGGCCATTGCGCAGG 2100

W Q A A G V R P D A V I G H S Q
GTGAGATCGCCGCGAGCTTGTGTGGCGGGTGCGGTGTCACGCGATGCC 2150
G E I A A A C V A G A V S L R D A
5 GCGCGGATCGTGACCTTCCGCGAGCCAGGCGATCGCCCGGGGCTGCGCGG 2200
A R E V T L R S Q A I A R G L A G
CGCGGGCGGATGSCATCCGTGCGCCCTGCCCGCGCAGGATGTGAGCTGG 2250
R G A M A S V A L P A Q D V E L
TCGACGGGGGCTGGATCGCCGCCACACGCGCGCCGCTCCACCTGATC 2300
V D G A W I A A H N G P A S T V I
10 GCGGGCACCCCGGAAGCGGTGACCATGTCTCACCCTCATGAGSCACA 2350
A G T P E A V D H V L T A H E A Q
AGGGGTGCGGGTGCGGGCGGATCACCGTCGACTATGCCTCGCACACCCCGC 2400
G V R V R R I T V D Y A S H T P
15 ACCTCGAGCTGATCCGCGACGAACACTCGACATCACTAGCGACAGCAGC 2450
H V E L I R D E L L D F T S D S S
TCGCGAGACCCCGCTCGTGCCGTGGCTGTGACCGTGGACGGCACCTGGGT 2500
S Q T P L V P W L S T V D G T W V
CGACAGCCCGCTGGACGGGGAGTACTGGTACCGGAACCTGCGTGAACCGG 2550
D S P L D G E Y W Y R N L R E P
20 TCGGTTTCCACCCCGCGCTCAGCCAGTTGCAGGCCAGGGCGACACCGTG 2600
V G F H P A V S Q L Q A Q G D T V
TTGCTCGAGGTCAGCGCCAGCCCGGTGTTGTTGAGGCGATGGACGACGA 2650
F V E V S A S P V L L Q A M E D D
TGTCGTCACGGTTCCACGCTGCGTGTGACGACGGCGACGCCACCCGGA 2700
25 V V T V A T L R R D D G D A T R
TGCTCACCGCCCTGGCACAGGCCTATGTCCACGGCGTCACCGTCGACTGG 2750
M L T A L A Q A Y V H G V T V D W
CCCGCCATCCTCGGACACCACCAACCCGGGTACTGGACCTTCCGACCTA 2800
P A I L G T T T T T R V L D L F T Y
30 CGCCTTCCAACACCAGCGGTACTGGCTCGAGTCGGCACGCCCGGCCGCAT 2850
A F Q H Q R Y W L E S A R P A A
CCGACGCGGGCCACCCCGTGTCTGGGCTCCGGTATCGCCCTCGCCGGGTG 2900
S D A G H P V L G S G I A L A G S
CCGGGCGGGGTGTTACGGGTTCGGTGCCGACCGGTGCGGACCGCGCGGT 2950
35 P G R V F T G S V P T G A D R A V
GTTCTGTCGCGAGCTGGCGCTGGCCCGCGGACGCGGTGACTGCGCCA 3000
F V A E L A L A A A D A V D C A
CGGTGAGCGGCTCGACATCGCCTCCGTGCCCGGCCGCGGGCCATGGC 3050
T V E R L D I A S V P G R P G H G
40 CGGACGACCGTACAGACCTGGGTGACGAGCCGGCGGACGACGGCCGGCG 3100
R T T V Q T W V D E P A D D G R R
CCGTTTACCGTGCACACCCGACCGGCGACGCCCGGTGGACGCTGCACG 3150
R F T V H T R T G D A P W T L H
CCGAGGGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATGCGGCCGAC 3200
45 A E G V L R P H G T A L P D A A D
GCCGAGTGGCCCCACCGGGCGCGGTGCCCGCGGACGGGCTGCCGGGTGT 3250
A E W P P P G A V P A D G L P G V
GTGGCGCCGGGGGACAGGTCTTCGCCGAGGCCGAGGTGGACGGACCGG 3300
W R R G D Q V F A E A E V D G P
50 ACGGTTTCTGTTGACCCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTC 3350
D G F V V H P D L L D A V F S A V
GGCGACGGAAGCCGCCAGCCGGCCGGATGGCGCGACCTGACGGTGCACGC 3400
G D G S R Q P A G W R D L T V H A
GTCGGACGCCACCGTACTGCGCGCTGCCTCACCGGGCGCACCGACGGAG 3450
55 S D A T V L R A C L T R R T D G
CCATGGGATTGCGCCGCTTCGACGGCGCCGCTGCCGGTACTCACCGCG 3500
A M G F A A F D G A G L P V L T A
GAGGCGGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTC 3550
E A V T L R E V A S P S G S E E S
60 GGACGCCCTGCACCGGTTGGAGTGGCTCGCGGTGCGCGAGGCGGTCTACG 3600

D G L H R L E W L A V A E A V Y
 ACGGTGACCTGCCCGAGGGACATGTCTGATCACCAGCCGCCACCCCGAC 3650
 D G D L P E G H V L I T A A H P D
 GACCCCGAGGACATACCCACCCCGCCGACCCCGCCGCGCTCCT 3700
 5 C P E D I P T R A H T R A T R V L
 GACCCGCTGCAACACCACCTCACCACCACCGACACACCCCTCATCGTCC 3750
 T A L Q H H L T T T D H T L I V
 ACACCACCACCGACCCCGCCGCGCCACCGTCACCGGCTCACCAGCACC 3800
 H T T T D P A G A T V T G L T R T
 10 CCCCAGAAACGAACCCCCACCGCATCCGCTCATCGAAACCGACACCC 3850
 A Q N E H P H R I R L I E T D H P
 CCACACCCCTCCCTGGCCCACTCGCCACCCTCGACCACCCCGAC 3900
 H T P L P L A Q L A T L D H P H
 TCCGCTCACCACACACCTCCACCACCCCGACCTCACCCTCCAC 3950
 15 L R L T H H T L H H P H L T P L H
 ACCACACCCCGACCCACACCCCTCAACCCCGAACACGCCATCAT 4000
 T T T P P T T T P L N P E H A I I
 CATACCGGCGCTCCGGCACCTCGCCGGCATCCTCGCCCGCACCTGA 4050
 I T G G S G T L A G I L A R H L
 20 ACCACCCCGACACCTACCTCCTCTCCCGACCCCGACCCCGACGCCACC 4100
 N H P H T Y L L S R T P P P D A T
 CCGGCGACCCACCTCCCTGCGACGTGCGCGACCCCGACCAACTCGCCAC 4150
 P G T H L P C D V G D P H Q L A T
 CACCTCACCACATCCCCCAACCCCTCACCAGCATCTTCCACACCGCG 4200
 25 T L T H I P Q P L T A I F H T A
 CCACCTCGACGACGGCATCCTCCACGCTCACCCTCGACCGCTCACC 4250
 A T L D D G I L H A L T P D R L T
 ACGTCTCCACCCCAAAGCCAACGCGCTGGCACCTGCACCACCTCAC 4300
 T V L H P K A N A A W H L H H L T
 30 CCAAAACCAACCCCTCACCCTTCTGCTCCTTACTCCAGCGCGCGCGCG 4350
 Q N Q P L T H F V L Y S S A A A
 TCCTCGGCGACCCCGGACAAGGAACTACGCGCGCGCAACGCTTCTC 4400
 V L G S P G Q G N Y A A A N A F L
 GACGCTCGCCACCCACCGCCACACCTCGGCAACCCGCGCACCTCCAT 4450
 35 D A L A T H R H T L G Q P A T S I
 CGCTGGGGCATGTGGCACACCAGCACCTCACCAGCAACTCGACG 4500
 A W G M W H T T S T L T G Q L D
 ACGCGCGCGGACCGCATCCGCGCGCGGTTTCTCCCGATCACGGAC 4550
 D A D R D R I R R G G F L P I T D
 40 GACGAGGGCATGGGGATGCAT
 D E G

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with
 the endogenous AT domain replaced by the AT domain of module 12 (specific for malonyl
 45 CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence
 shown below.

AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGCGAGAGCACC 50
 Q L A E A L L T L V R E S T
 GCGCGCTGCTCGCCACGTGGGTGGCGAGGACATCCCGCGACGGCGGC 100
 50 A A V L G H V G G E D I P A T A A
 GTTCAAGGACCTCGGCATCGACTCGCTCACCAGCGGTCCAGCTGCGCAACG 150
 F K D L G I D S L T A V Q L R N
 CCCTCACCAGGCGACCGGTGTGCGGCTGAACGCCACGGCGGTCTTCGAC 200
 A L T E A T G V R L N A T A V F D
 55 TTCCCGACCCCGACGTGCTCGCGGGAAGCTCGGCGACGAAGTACCGG 250
 F P T P H V L A G K L G D E L T G
 CACCCGCGCGCGCTGCTGCCCCGACCGCGCGCCACGGCGGTGCGCACG 300

T R A F V V P R T A A T A G A H
ACGAGCCGCTGGCGATCGTGGGAATGGCCTGCCGGCTGCCCGGCGGGGTC 350
D E P L A I V G M A C R L P G G V
GGGTCAACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
5 A S P E E L W H L V A S G T D A I
CACGGAGTTCCCGACGGACCGCGGCTGGGACGTCGACGCGATCTACGACC 450
T E F F T D R G W D V D A I Y D
CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
F D P D A I G K T F V R H G G F L
10 ACCGGCGCGACAGGCTTCGACGCGGCGTTCTTCGGCATCAGCCCGCGCGA 550
T G A T G F D A A F F G I S P R E
GGCCCTCGCGATGGACCCGCGAGCAGCGGGTGCTCCTGGAGACGTCGTGGG 600
A L A M D P Q R V L L E T S W
AGGCGTTCGAAAGCGCGGCATCACCCCGGACTCGACCCGCGGCAGCGAC 650
15 E A F E S A G I T P D S T R G S D
ACCGGCGTGTTTCGTGGCGCCTTCTCCTACGGTTACGGCACCGGTGCGGA 700
T G V F V G A F S Y G Y G T G A D
CACCGACGGCTTCGGCGCGACCGGCTCGCAGACAGTGTGCTCTCCGGCC 750
T D G F G A T G S Q T S V L S G
20 GGCTGTCGTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800
R L S Y F Y G L E G P A V T V D T
CGGTGTTTCGTGCTGCTGGTGGCGCTGCACCAGGCCGGGCAGTCGCTGCG 850
A C S S S L V A L H Q A G Q S L R
CTCCGGCGGAATGCTCGCTCGCCCTGGTGGCGGCGTACGGTGATGGCGT 900
25 S G E C S L A L V G G V T V M A
CTCCCGGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950
S P G F V E F S R Q R G L A P D
GGCCGGGCGAAGGCGTTCCGGCGCGGTGCGGACGGCACGAGCTTCGCCGA 1000
G R A K A F G A G A D G T S F A E
30 GGGTGCCGGTGTGCTGATCGTCGAGAGGCTCTCCGACGCCGAACGCAACG 1050
G A G V L I V E R L S D A E R N
GTCACACCGTCCTGGCGGTGCTCCGTGGTTCGGCGGTCAACCAGGATGGT 1100
G H T V L A V V R G S A V N Q D G
GCCTCCAACGGGCTGTGGCGCCGAACGGGCGTTCGAGGAGCGGGTGAT 1150
35 A S N G L S A P N G P S Q E R V I
CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGCGGACGTGGACGCCG 1200
R Q A L A N A G L T P A D V D A
TCGAGGCCCCACGGCACCGGCACCAGGCTGGGCGACCCCATCGAGGCACAG 1250
V E A H G T G T R L G D P I E A Q
40 GCGGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCCCTGCTGCTGGG 1300
A V L A T Y G E R A T P L L L G
CTCGCTGAAGTCCAACATCGGGCACGCCCAGGCCGCTCCGGCGTCCGCCG 1350
S L K S N I G H A Q A A S G V A
GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGGAGCTGCCGCCGACG 1400
45 S I I K M V Q A L R H G E L P P T
CTGCACGCCGACGAGCCGTGCGCGCACGTGCGACTGGACGGCCGGCGCCGT 1450
L H A D E P S P H V D W T A G A V
CGAFACTGCTGACGTCCGGCCCGGCGTGGCCCCGAGACCGACCGGCCACGGC 1500
E L L T S A R P W P E T D R P R
50 GTGCCGCCGTCTCCTCGTTCCGGGTGAGCGGCACCAACGCCACGTCATC 1550
R A A V S S F G V S G T N A H V I
CTGSAGGCCGGACCGGTAACGGAGACGCCCCGCGGCATCGCCTTCGGGTGA 1600
L E A G P V T E T P A A S P S G D
CCTTCCCTGCTGCTGTCGGCACGCTCACCGGAAGCGCTCGACGAGCAGA 1650
55 L P L V S A R S P E A L D E Q
TCTGCCGACTGCGCGCCTACCTGGACACACCCCGGACGTGACCGGGTG 1700
T R R L R A Y L D T T P D V D R V
GCGGTGGCAGACGCTGGCCCCGGCGCACACACTTCGCCCCACCGCGCGT 1750
A V A Q T L A R R T H F A H R A V
60 GCTGCTCGGTGACACCGTCATCACACACCCCCCGCGGACCGGCCCGGACG 1800

SUBSTITUTE SHEET (RULE 26)

L L G E T V I T T P P A D R P D
 AACTCGTCTTCTCTACTCCGGCCAGGGCAGCCATCCCGCGATGGGC 1850
 E L V F V Y S G Q G T Q H P A M G
 GAGCAGCTAGCGGCGCGCTTCCCGCTCTTCGCGCGGATCCATCAGCAGGT 1900
 5 E Q L A A A F P V F A E I H Q Q V
 GTGGGACCTGCTCGATGTGCCCGATCTGGAGGTGAACGAGACCGGTTACG 1950
 W D L L D V P D L E V N E T G Y
 CCCAGCCGGCCCTGTTCGCAATGCAGGTGGCTCTGTTCCGGGCTGCTGGAA 2000
 A Q P A L F A M Q V A L F G L L E
 10 TCGTGGGCTGTACGACCGGACGCGGTGATCGGCCATTCGGTGGGTGAGCT 2050
 S W G V R P D A V I G H S V E L
 TGCGGCTCCGTATGTGTCCGGGTGTGGTCTGTGGAGSATGCCTGCACCT 2100
 A A A Y V S G V W S L E D A C T
 TGGTGTCCGGCGCGGCTCCTCTGATGCAGGCTCTCCCGCGGCTGGGGTG 2150
 15 L V S A R A R L M Q A L F A S G V
 ATGGTCCCTGTCCCGGTCTCGGAGGATGAGGCCCGGCGCGGTGCTGGGTGA 2200
 M V A V P V A E D E A R A V L G E
 GGGTGTGAGATCGCCCGGTCACCGCCCGTCTCGGTGGTTCTCTCCG 2250
 G V E I A A V N G P S S V V L S
 20 GTGATGAGGCCCGCGTGTGTCAGGCCGCGGAGGGGCTGGGGAAGTGGACG 2300
 G D E A A V L Q A A E G L G K W T
 CGGCTGGCGACCCAGCCACGCGTTCATTCCGCCCCGTATGGAACCCATGCT 2350
 R L A T S H A F H S A R M E F M L
 GGAGGAGTCCCGGCGGTCCCGAAGGCCTGACCTACCGGACCGCCGAGG 2400
 25 E E F R A V A E G L T Y R T P Q
 TCTCCATGGCCGTTGGTGTATCAGGTGACCACCGCTGAGTACTGGGTGCGG 2450
 V S M A V G D Q V T T A E Y W V R
 CAGGTCCGGGACACGGTCCGGTTCGGCGAGCAGGTGGCCTCGTACGAGGA 2500
 Q V R D T V R F G E Q V A S Y E D
 30 CGCCGTGTTCTGTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCCTGGTTCG 2550
 A V F V E L G A D R S L A R L V
 ACGGTGTCCGCGATGCTGCACGGCGACCACGAAATCCAGGCCGCGATCGGC 2600
 D G V A M L H G D H E I Q A A I G
 GCCCTGGCCACCTGTATGTCAACGGCGTCACGGTCCACTGGCCCGCGCT 2650
 35 A L A H L Y V N G V T V D W F A L
 CCTGGGCGATGCTCCGGCAACACGGGTGCTGGACCTTCGACATACGCT 2700
 L G D A P A T R V L D L P T Y A
 TCCAGCACCAGCGCTACTGGCTCGAGTCCGGCACGCCCGCGCATCCGAC 2750
 F Q H Q R Y W L E S A R P A A S D
 40 GCGGGCCACCCCGTGTGGGTCCGGTATCGCCCTCGCCGGGTGCGCGGG 2800
 A G H P V L G S G I A L A G S P G
 CCGGGTGTTCACGGGTCCGTGCCGACCGGTGCGGACCGCGCGGTGTTCG 2850
 R V F T G S V P T G A C R A V F
 TCGCCGAGCTGGCGCTGGCCCGCGGACGCGGTCCACTGCGCCACGGTC 2900
 45 V A E L A L A A A D A V D C A T V
 GAGCGGCTCGACATCGCCTCCGTGCCCGGCCGCGCGGCGCATGGCCGGAC 2950
 E R L D I A S V P G R P G H G R T
 GACCGTACAGACCTGGGTGACGAGCCGGCGGACGACGGCCGCGCGCGGT 3000
 T V Q T W V D E P A D C G R R R
 50 TCACCGTGCACACCCGACCCGGCGACGCCCCGTGGACGCTGCACGCCGAG 3050
 F T V H T R T G D A P W T L H A E
 GGGGTGCTGCGCCCCCATGGCACGGCCCTGCCCGATCGGGCCGACGCCGA 3100
 G V L R F H G T A L P D A A C A E
 55 GTGGCCCCACCGGCGCGGTGCCCGCGGACGGGTCCCGGGTGTGTGGC 3150
 W P P P G A V P A D G L P G V W
 GCCGGGGGGACAGGTCTTCGCGGAGGCCGAGGTGGACGGACCGGACGGT 3200
 R R G D Q V F A E A E V C G F D G
 TTCGTGCTGACCCCGACCTGCTGACGCGGTCTTCTCCGCGGTGCGCGA 3250
 F V V H P D L L D A V F S A V G D
 60 CGGAAGCCGCCAGCCGGCCGGATGGCGGACCTGACGGTGCACGCGTCCG 3300

SUBSTITUTE SHEET (RULE 26)

D S R Q P A G W R D L T V H A S
 ACGCCACCGTACTGCGCCCTGCCTCACCCGGCGCACCGACGGAGCCATG 3350
 D A T V L R A C L T R R T D G A M
 5 GATTCCCGGCTTCGACGGCGCGGCTGCGGTAATCACCCGGAGGC 3400
 G F A A F D G A G L P V L T A E A
 GGTGACGCTGCGGGAGGTGGCGTCACCGTCCGGCTCCGAGGAGTCGGACG 3450
 V T L R E V A S P S G S E E S D
 GCGTCACCGGTTGGAGTGGCTCGCGGTGCGGAGGCGGTCTACGACGGT 3500
 G L H R L E W L A V A E A V Y D G
 10 GACCTGCCCCGAGGGACATGTCCTGATCACCGCGGCCACCCCGACGACCC 3550
 D L P E G H V L I T A A H P D D P
 CGAGGACATACCCACCCGCGCCACACCGCGCCACCCGCGTCTTGACCG 3600
 E D I P T R A H T R A T R V L T
 CGCTGCAACACCACCTCACCAACCGACACACCTCATCGTCCACACC 3650
 15 A L Q H H L T T T D H T L I V H T
 ACCACCGACCCCGCGGCGCCACCGTCACCGGCTCACCCGACCCGCCA 3700
 T T D P A G A T V T G L T R T A Q
 GAACGAACACCCCAACCGCATCCGCCTCATCGAAACCGACACCCCCACA 3750
 N E H P H R I R L I E T D H P H
 20 CCGCCCTCCCGTGGCCCACTCGCCACCCCTCGACACCCCCACCTCCGC 3800
 T P L P L A Q L A T L D H P H L R
 CTCACCCACCAACCCCTCCACCAACCCCACTCACCCCCCTCCACACCAC 3850
 L T H H T L H H P H L T P L H T T
 CACCCACCCACCAACCCCCCTCAACCCCGAACACGCCATCATCATCA 3900
 25 T P P T T T P L N P E H A I I I
 CCGGCGGCTCCGGCACCCCTCGCGGCGATCCTCGCCCGCCACCTGAACCAC 3950
 T G G T L A G I L A R H L N H
 CCCCACACCTACCTCCTCTCCCGCACCCCAACCCCGACGCCACCCCGG 4000
 P H T Y L L S R T P P P D A T P G
 30 CACCCACCTCCCGTGGGACGTGGGCGACCCCAACCACTCGCCACCAACC 4050
 T H L P C D V G D P H Q L A T T
 TCACCCACATCCCCCAACCCCTCACCGCCATCTTCCACACCGCGGCCACC 4100
 L T H I P Q P L T A I F H T A A T
 35 CTGACGACGGCATCCTCCACGCCCTCACCCCGACCGCCTCACCAACCGT 4150
 L D D G I L H A L T P D R L T T V
 CCTCCACCCCAAGCCAAAGCGCGCTGGCACCTGCACCACTCACCCAAA 4200
 L H P K A N A A W H L H H L T Q
 ACCAACCCCTCACCACTTCGTCTCTACTCCAGCGCGCGCGCGTCTCTC 4250
 N Q P L T H F V L Y S S A A A V L
 40 GGCAGCCCGGACAAGGAACTACGCCCGCGCAACGCCTTCCTCGACGC 4300
 G S P G Q G N Y A A A N A F L D A
 CCTCGCCACCCACCGCCACACCCCTCGGCCAACCGCCACCTCCATCGCCT 4350
 L A T H R H T L G Q P A T S I A
 GGGGCATGTGGCACACCACGACCCCTCACCGGACAACCTCGACGACGCC 4400
 45 W G M W H T T S T L T G Q L D D A
 GACCGGGACCGCATCCGCGCGGCGGTTTCCTCCCGATCACGGACGACGA 4450
 D R D I R R G G F L P I T D D E
 GGGCATGGGGATGCAT
 G

50

The *NheII-XhoI* restriction fragment that encodes module 8 of the FK-520 PKS with the endogenous AT domain replaced by the AT domain of module 13 (specific for methylmalonyl CoA) of the rapamycin PKS has the DNA sequence and encodes the amino acid sequence shown below.

55 AGATCTGGCAGCTCGCCGAAGCGCTGCTGACGCTCGTCCGGGAGAGCACC 50
 Q L A E A L L T L V R E S T
 GCCGCCGTGCTCGGCCACGTGGGTGGCGAGGACATCCCGCGACGGCGGC 100

A A V L G H V G G E D I P A T A A
 GTTCAAGGACCTCGGCATCGACTCGCTCAGCCCGGTCCAGCTGCGCAACG 150
 F K D L G I D S L T A V Q L R N
 5 GTTCAAGGACCTCGGCATCGACTCGCTCAGCCCGGTCCAGCTGCGCAACG 200
 A L T E A T G V R L N A T A V F D
 TTCCCGACCCCGCACGTGCTCGCCGGGAAGCTCGGCGACGAAGTACCGG 250
 F P T P H V L A G K L G D E L T G
 CACCCGCGCGCCCGTCTGTGCCCCGGACCGCGGCCACGGCCGGTGCACG 300
 T R A P V V P R T A A T A G A H
 10 ACGAGCCGCTGCGCATCGTGGGAATGGCCTGCGGCTGCCCCGGCGGGGTC 350
 D E P L A I V G M A C R L P G G V
 GCGTCACCCGAGGAGCTGTGGCACCTCGTGGCATCCGGCACCGACGCCAT 400
 A S P E E L W H L V A S G T D A I
 CACGGAGTTCCCGACGGACCGCGGCTGGGACGTGACGCGATCTACGACC 450
 15 T E F P T D R G W D V D A I Y D
 CGGACCCCGACGCGATCGGCAAGACCTTCGTCCGGCACGGTGGCTTCCTC 500
 P D P D A I G K T F V R H G G F L
 ACCGCGCGCAGGCTTCGACGCGGCTTCTTCGGCATCAGCCCGCGCGA 550
 T G A T G F D A A F F G I S P R E
 20 GGGCCCTCGGATGGACCCCGCAGCAGCGGGTGTCTCTGGAGACGTCTGGG 600
 A L A M D P Q Q R V L L E T S W
 AGSCGTTTCAAAGCCCGGCATCAGCCCGGACTCGACCCCGCGCAGCGAC 650
 E A F E S A G I T P D S T R G S D
 ACCGCGGTGTTCTGTCGGCGCCTTCTCTACGGTTACGGCACCGGTGCGGA 700
 25 T G V F V G A F S Y G Y G T G A D
 CACCGACGGCTTCGGCGCAGCGGCTCGCAGACCAAGTGTGCTCTCCGGCC 750
 T D G F G A T G S Q T S V L S G
 GGCTGTCTACTTCTACGGTCTGGAGGGTCCGGCGGTACGGTTCGACACG 800
 R L S Y F Y G L E G P A V T V D T
 30 GCGTGTCTGTCGCTCGCTGGTGGCGCTGCACCAGGCCGGGCAGTCTGCTGCG 850
 A C S S L V A L H Q A G Q S L R
 CTCCGCGCAATGCTCGCTCGCCCTGGTCCGGCGCGCTCACGGTGATGGCGT 900
 S G E C S L A L V G G V T V M A
 CTCCCGCGCGGCTTCGTGGAGTTCTCCCGGCAGCGCGGCTCGCGCCGGAC 950
 35 S P G G F V E F S R Q R G L A P D
 GGCCGGGCGAAGGCGTTCGGCGCGGGTGCAGGACGGCACGAGCTTCGCCGA 1000
 G R A K A F G A G A D G T S F A E
 GGGTGCCGCTGTGCTGATCGTCTGAGAGGCTCTCCGACGCCGAACGCAACG 1050
 G A G V L I V E R L S D A E R N
 40 GTCAACCGCTCTGGCGGTCTGCTGGTTCGGCGGTCAACAGGATGGT 1100
 G H T V L A V V R G S A V N Q D G
 GCCTCCAACGGGCTGTGCGCGCCGAACGGGCGGTTCGACAGGAGCGGGTGAT 1150
 A S N G L S A P N G P S Q E R V I
 CCGGCAGGCCCTGGCCAACGCCGGGCTCACCCCGCGCGACGTGGACGCCG 1200
 45 R Q A L A N A G L T P A D V D A
 TCGAGGCCACGGCACCGGCACAGGCTGGGCGACCCCATCGAGGCACAG 1250
 V E A H G T G T R L G D P I E A Q
 GCGTACTGGCCACCTACGGACAGGAGCGCGCCACCCCGCTGCTGCTGGG 1300
 A V L A T Y G Q E R A T P L L L G
 50 CTCGCTGAAGTCCAACATCGGCCACGCCAGGCCGCGTCCGGCGTCCGCCG 1350
 S L K S N I G H A Q A A S G V A
 GCATCATCAAGATGGTGCAGGCCCTCCGGCACGGGAGCTGCCGCCGACG 1400
 G I I K M V Q A L R H G E L P P T
 CTGCACGCCGACGAGCCGTGCGCGCACGTGCGTGGACGGCCGGCGCCGT 1450
 55 L H A D E P S P H V D W T A G A V
 CGAAGTGTGACGTCCGGCCCGCGCGTGGCCCGAGACCGACCGGCCACGGC 1500
 E L L T S A P P W P E T D P P R
 GTCCCGCGCTCTGCTGCTGCGGCTGAGCGGCTACCAACGCCCTACGTCATC 1550
 R A A V S F G V S G T N A H V I
 60 CTCGAGGCCGACCGGTAAACGGAGACGCCCGCGGCATCGCCTTCGGGTGA 1600

SUBSTITUTE SHEET (RULE 26)

D E F A D D G R R R F T V H T R T
CGGCGACGCGCGCTGGACGCTGCACGCCGAGGGGGTGTCTGCGCGCGCGATG 3150
G D A F W T L H A E G V L P P H
5 D G C G G C T T G G A T G G G G G A C G C G A G T G G G G G A C C G G G G 3200
S T A L F D A A D A E W P P P G A
GTGCGCGCGCGCGCTGCGGGGTGTGTGGCGCGCGGGGGACCGGTCTT 3250
V P A E G L P G V W R R G D Q V F
CGCGGAGGCGCGAGGTGGACGGACCGGACGGTTTTCGTGGTGCACCCCGACC 3300
A E A E V D G P D G F V V H P D
10 TGCTCGACGCGGTCTTCTCCGCGGTGCGCGACGGAAGCCGCCAGCCGGCC 3350
L L D A V F S A V G D G S R Q P A
GGATGGCGCGAGCTGACGGTGCACGGTCCGACGCCACCGTACTGCGCGC 3400
G W R D L T V H A S D A T V L R A
CTGCCCTACCCCGCGCACCGACGGAGCCATGGGATTCGCCGCTTCGACG 3450
15 C L T R R T D G A M G F A A F D
CGCGCGCGCTGCGGTACTCACCGCGGAGGCGGTGACGCTGCGGGAGGTG 3500
G A G L P V L T A E A V T L R E V
GCGTCACCGTCCCGCTCCGAGGAGTCGGACGGCCTGCACCGGTGGAGTG 3550
A S P S G S E E S D G L H R L E W
20 GCTCGCGGTGCGCGAGGCGGTCTACGACGGTGACCTGCGCGAGGGACATG 3600
L A V A E A V Y D G D L P E G H
TGCTGATCAGCGCGCGCGCGCGACGACCGCGAGGACATACCGCGCGCG 3650
V L I T A A H P D D P E D I P T R
GCCACACCGCGCGCGCGCGCGCTCTGACCGCGCTGCAACACCCACCTCAC 3700
25 A H T R A T R V L T A L Q H H L T
CACCACCGACCGACCGCTCATCGTCCACACCGACCGCGCGCGCGCG 3750
T T D H T L I V H T T T D P A G
CCACCGTCCACCGCGCTCACCGCGACCGCGCGAGAACGAACACCGCGCG 3800
A T V T G L T R T A Q N E H P H R
30 ATCCGCTCATCGAAACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3850
I R L I E T D H P H T P L P L A Q
ACTCGCGACCGCTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3900
L A T L D H P H L R L T H H T L
ACCACCG 3950
35 H H F H L T P L H T T T P P T T T
CCCGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4000
P L N P E H A I I I T G G S G T L
CG 4050
A G I L A R H L N H P H T Y L L
40 CG 4100
S R T P P P D A T P G T H L P C D
GTGCG 4150
V G C F H Q L A T T L T H I P Q P
CCTCACCG 4200
45 L T A I F H T A A T L D D G I L
ACG 4250
H A L T P D R L T T V L H P K A N
GCCG 4300
A A W H L H H L T Q N Q P L T H F
50 CGTCCTCTACTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4350
V L Y S S A A A V L G S P G Q G
ACTACCG 4400
N Y A A A A F L D A L A T H R H
ACCG 4450
55 T L S Q P A T S I A W G M W H T T
CAGCACCG 4500
S T L T G Q L D D A D R D P I R
CG
R G G F L P I T D D E G

SUBSTITUTE SHEET (RULE 26)

Phage KC515 DNA was prepared using the procedure described in Genetic Manipulation of *Streptomyces*. A Laboratory Manual, edited by D. Hopwood *et al.* A phage suspension prepared from 10 plates (100 mm) of confluent plaques of KC515 on *S. lividans* TK24 generally gave about 3 µg of phage DNA. The DNA was ligated to circularize at the cos site, subsequently digested with restriction enzymes *Bam*HI and *Pst*I, and dephosphorylated with SAP.

Each module 8 cassette described above was excised with restriction enzymes *Bg*II and *Nsi*I and ligated into the compatible *Bam*HI and *Pst*I sites of KC515 phage DNA prepared as described above. The ligation mixture containing KC515 and various cassettes was transfected into protoplasts of *Streptomyces lividans* TK24 using the procedure described in Genetic Manipulation of *Streptomyces*, A Laboratory Manual edited by D. Hopwood *et al.* and overlaid with TK24 spores. After 16-24 hr, the plaques were restreaked on plates overlaid with TK24 spores. Single plaques were picked and resuspended in 200 µL of nutrient broth. Phage DNA was prepared by the boiling method (Hopwood *et al.*, *supra*). The PCR with primers spanning the left and right boundaries of the recombinant phage was used to verify the correct phage had been isolated. In most cases, at least 80% of the plaques contained the expected insert. To confirm the presence of the resistance marker (thiostrepton), a spot test is used, as described in Lomovskaya *et al.* (1997), in which a plate with spots of phage is overlaid with mixture of spores of TK24 and phiC31 TK24 lysogen. After overnight incubation, the plate is overlaid with antibiotic in soft agar. A working stock is made of all phage containing desired constructs.

Streptomyces hygroscopicus ATCC 14891 (see US Patent No. 3,244,592, issued 5 Apr 1966, incorporated herein by reference) mycelia were infected with the recombinant phage by mixing the spores and phage (1×10^8 of each), and incubating on R2YE agar (Genetic Manipulation of *Streptomyces*, A Laboratory Manual, edited by D. Hopwood *et al.*) at 30°C for 10 days. Recombinant clones were selected and plated on minimal medium containing thiostrepton (50 µg/ml) to select for the thiostrepton resistance-conferring gene. Primary thiostrepton resistant clones were isolated and purified through a second round of single colony isolation, as necessary. To obtain thiostrepton-sensitive revertants that underwent a second recombination event to evict the phage genome, primary recombinants were propagated in liquid media for two to three days in the absence of thiostrepton and then spread on agar medium without thiostrepton to obtain spores. Spores were plated to obtain about 50 colonies per plate, and thiostrepton sensitive colonies were identified by

replica plating onto thiostrepton containing agar medium. The PCR was used to determine which of the thiostrepton sensitive colonies reverted to the wild type (reversal of the initial integration event), and which contain the desired AT swap at module 8 in the ATCC 14891-derived cells. The PCR primers used amplified either the KS/AT junction or the AT/DH junction of the wild-type and the desired recombinant strains. Fermentation of the recombinant strains, followed by isolation of the metabolites and analysis by LCMS, and NMR is used to characterize the novel polyketide compounds.

Example 2

10 Replacement of Methoxyl with Hydrogen or Methyl at C-13 of FK-506

The present invention also provides the 13-desmethoxy derivatives of FK-506 and the novel PKS enzymes that produce them. A variety of *Streptomyces* strains that produce FK-506 are known in the art, including *S. tsukubaensis* No. 9993 (FERM BP-927), described in U.S. Patent No. 5,624,852, incorporated herein by reference; *S. hygroscopicus* subsp. *yakushimaensis* No. 7238, described in U.S. patent No. 4,894,366, incorporated herein by reference; *S. sp.* MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference; and *S. sp.* MA 6548, described in Motamedi *et al.*, 1998, "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK-506," *Eur. J. Biochem.* 256: 528-534, and Motamedi *et al.*, 1997, "Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK-506," *Eur. J. Biochem.* 244: 74-80, each of which is incorporated herein by reference.

The complete sequence of the FK-506 gene cluster from *Streptomyces* sp. MA6548 is known, and the sequences of the corresponding gene clusters from other FK-506-producing organisms is highly homologous thereto. The novel FK-506 recombinant gene clusters of the present invention differ from the naturally occurring gene clusters in that the AT domain of module 8 of the naturally occurring PKSs is replaced by an AT domain specific for malonyl CoA or methylmalonyl CoA. These AT domain replacements are made at the DNA level, following the methodology described in Example 1.

30 The naturally occurring module 8 sequence for the MA6548 strain is shown below,
followed by the illustrative hybrid module 8 sequences for the MA6548 strains.

GCATCGCGCTGTACGAGGGGBCACGGCGCACCGGAAGTCCCGTGGTGCTG 50
M R L Y E A A R R T G S P V V V
GCGGCCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
35 A A A L D D A P D V P L L R G L R

CCGTACGACCGCTCCGGCGTGCCGCGCTCCGGGAACGCTCTCTCGCCGACC 150
R T T V R R A A V R E R S L A D
GCTCGCGGTGCTGCGCGACGAGCGCGCGGACGCGCTCCCTCGCGTTCCG 200
R S P T C P T T S A P T P P S R S
5 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCGGAAGACAT 250
S W N S T A T V L G H L G A E D I
CCCGGCGAAGACGCTTCAAGGAACCTCGGCATCGACTCGCTCACC CGG 300
P A T T T F K E L G I D S L T A
TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350
10 V Q L R N A L T T A T G V R L N A
ACAGCGGTCTTCGACTTTCCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400
T A V F D F P T P R A L A A R L G
CGACGAGCTGGCCGGTACCGCGCGCGCGCTCGCGGCCCGGACCGCGGCCA 450
D E L A G T R A P V A A R T A A
15 CCGCGGCGCGCGACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
T A A A H D E P L A I V G M A C R
CTGCCGGGCGGGGTGCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
L P G G V A S P Q E L W R L V A S
CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
20 G T D A I T E F P A D R G W D V
ACGCGCTCTACGACCGGACCGCGATCGGCAAGACCTTCGTCCGG 650
C A L Y D P D P D A I G K T F V R
CACGGCGGGCTTCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
H G G F L D G A T G F D A A F F G
25 GATCAGCCCCGCGCGAGGCCCTGGCCATGGACCCGACGAACGGGTGCTCC 750
I S P R E A L A M D P Q Q R V L
TGGAGACGTCTGGGAGGCGTTCGAAAGCGCGGCATCACCCCGGACGCG 800
L E T S W E A F E S A G I T P D A
GCGCGGGGCGAGCGACACCGGCGTGTTCATCGGCGCGTTCTCCTACGGGTA 850
30 A R G S D T G V F I G A F S Y G Y
CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTTCGACAGCA 900
G T G A D T N G F G A T G S Q T
GCGTGCTCTCCGGCCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
S V L S G R L S Y F Y G L E G P S
35 GTCACGGTTCGACACCGCTGCTCGTCACTGGTTCGCCCTGCACCAGGC 1000
V T V D T A C S S S L V A L H Q A
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050
G Q S L R S G E C S L A L V G G
TCACGGTGATGGCGTTCGCCCCGCGGATTCGTGAGTTCTCCCGGCAGCGC 1100
40 V T V M A S P G G F V E F S R Q R
GGGCTCGCGCGGACGGGCGGGCGAAGGCGTTCGCGCGGGGCGCGGACGG 1150
G L A P C G R A K A F G A G A D G
TACGAGCTTCGCGGAGGGCGCGGGTGCCTGGTGGTTCGAGCGGCTCTCCG 1200
T S F A E G A G A L V V E R L S
45 ACGCGGAGCGCCACGGCCACACCGTCTCGCCCTCGTACGCGGCTCCGCG 1250
D A E R H G H T V L A L V R G S A
GCTAACTCCGACGGCGCGTTCGAACGGTCTGTTCGGCGCGGAAACGGCCCTC 1300
A N S D G A S N G L S A P N G P S
CCAGGAACGCGTCCACACCGCCCTCGCGAACGCGAAACTCACCCCGG 1350
50 Q E R V I H Q A L A N A K L T P
CCGATGTTCGACGCGGTTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400
A D V D A V E A H G T G T R L G D
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
P I E A Q A L L A T Y G Q D R A T
55 GCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCGG 1500
P L L L G S L K S N I G H A Q A
TCTCAGGGGTTCGCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
A S G V A G I I K M V Q A I R H G
GAACTGCCGCGGACACTCCACGCGGACGAGCCGTTCGCCGACGTGACTG 1600
60 E L P P T L H A D E P S P H V D W

SUBSTITUTE SHEET (RULE 26)

GACGCCCGGTGCCCTCGAGCTCCTGACGTCGGCCCGGCCCTGGCCGGGGA 1650
T A G A V E L L T S A R P W P G
CCGGTCCGCCCGCGCCCGCTGCCGTCTCGTCTCGTTCGGCGTGGAGCGGCACG 1700
T G R F R R A A V S S F G V S G T
5 AACGCCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCTGA 1750
N A H I I L E A G P V K T G P V E
GGCAGGAGCGATCGAGGCAGGACCGGTCTGAAGTAGGACCGGTCTGAGGCTG 1800
A S A I E A G P V E V G P V E A
10 GACCCCTCCCGCGCGGCCCGCTCAGCACCGGCGGAGACTTTCCGCTG 1850
G P L P A A P P S A P G E D L P L
CTCGTCTCGGCCGCTTCCCGGAGGCACTCGACGAGCAGATCGGGCGCCT 1900
L V S A R S P E A L D E Q I G R L
GCGCGCCTATCTCGACACCGGCCCGGGCGTCTGACCGGGCGGCGCTGGCGC 1950
R A Y L D T G P G V D R A A V A
15 AGACACTGGCCCGCGGTACGCACTTCACCCACCGGGCCGTACTGCTCGGG 2000
Q T L A R R T H F T H R A V L L G
GACACCGTTCATCGGCGCTCCCCCGCGGACCAGGCCGACGAACCTCGTCTT 2050
D T V I G A P A D Q A D E L V F
CGTCTACTCCGGTCAAGGCACCCAGCATCCCGCGATGGGCGAGCAACTCG 2100
20 V Y S G Q G T Q H P A M G E Q L
CGGCCGCGTTCGCCGTGTTCCCGCATGCCTGGCACGACGCGCTCCGACGG 2150
A A A F F V F A D A W H D A L R R
CTCGACGACCCCGACCCGACGACCCACACGGAGCCAGCACACGCTCTT 2200
L D D P D P H D P T R S Q H T L F
25 CGCCACCGGCGGCTTCACCGCCCTCCTGAGGTCTCTGGGACATCACGC 2250
A H Q A A F T A L L R S W D I T
CGCACGCGTTCATCGGCCACTCGCTCGGCGAGATCACCGCCGCGTACGCC 2300
P H A V I G H S L G E I T A A Y A
30 GCGGGATCCTGTCTGCTCGACGACGCTGCACCCCTGATCACACGCGTGC 2350
A G I L S L D D A C T L I T T R A
CCGCCTCATGCACACGCTTCCGCCGCGCGGCCATGGTTCACCGTGTCTGA 2400
R L M H T L P P P G A M V T V L
CCAGCGAGGAGGAGGCCCGTCAAGGCGCTGCGGCCGGGCGTGGAGATCGCC 2450
T S E E E A R Q A L R P G V E I A
35 GCGGTCTTCGGCCCGCACTCCGTCTGTCTCTCGGGCGACGAGGACGCCGT 2500
A V F G P H S V V L S G D E D A V
GCTCGACGTCTGCACAGCGGCTCGGCATCCACCACCGTCTGCCCGCGCCGC 2550
L D V A Q R L G I H H R L P A P
ACGCGGGCACTCCGCGCATGGAACCCGTGGCCGCGAGCTGCTCGCC 2600
40 H A G H S A H M E P V A A E L L A
ACCACTCGCGAGCTCCGTTACGACCGGCCCCACACCGCCATCCCGAACGA 2650
T T R E L R Y D R P H T A I P N D
CCCCACCAACCGCGAGTACTGGGCCGAGCAGGTCCGCAACCCCGTGTGT 2700
P T T A E Y W A E Q V R N P V L
45 TCCACGCCCCACCCAGCGGTACCCCGACGCGGTCTTCGTGAGATCGGC 2750
F H A H T Q R Y P D A V F V E I G
CCCGGCCAGGACCTCTCACCGCTGGTCTGACGGCATCGCCCTGAGAACGG 2800
P G Q D L S P L V D G I A L Q N G
50 CACGGCGGACGAGGTGCACGCGCTGCACACCGCGCTCGCCCGCCTCTTCA 2850
T A D E V H A L H T A L A R L F
CACGCGGCGCCACGCTCGACTGGTCCCGCATCCTCGGCGGTGCTTCGCGG 2900
T R G A T L D W S R I L G G A S R
CACGACCCCTGACGTCCCTCTGACGCTTCCAGCGGCGTCCCTACTGGAT 2950
H D P D V P S Y A F Q R R P Y W I
55 CGAGTCCGCTCCCCGSCACGGCCGACTCGGGCCACCCCGTCTCTCGGCA 3000
E S A P P A T A D S G H P V L G
CCGGAGTCCCGCTCCCGGGTCCCGGGCGGGGTGTTACGGGTCCCGT 3050
T G V A V A G S P G R V F I G P V
CCCGCCGGTGGGACCGCGCGGTGTTTCATCGCCGAACTGGCGCTCGCCGC 3100
60 P A G A D R A V F I A E L A L A A

SUBSTITUTE SHEET (RULE 26)

CGCCGACGCCACCGACTGCGCCACGGTCTGAACAGCTCGACGTCACCTCCG 3150
A D A T D C A T V E Q L D V T S
TGCCCGGCGGATCCGCCCGCGGCAGGGCCACCGCGCAGACCTGGGTGAT 3200
P G G S A R G R A T A T W V D
5 GAACCCGCGCGCGACGGGCGGGCGCGCTTCACCGTCCACACCCGCGTCGG 3250
E P A A D G R R R F T V H T R V G
CGACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCCGGCGCG 3300
D A P W T L H A E G V L R P G R
TGCCCCAGCCGAAGCCGTGCACACCGCCTGGCCCCCGCGGGCGCGGTG 3350
10 V P Q P E A V D T A W P P P G A V
CCCGCGGACGGGCTGCCCGGGGCGTGGCGACGCGCGGACCAGGTCTTCGT 3400
P A D G L P G A W R R A D Q V F V
CGAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGC 3450
E A E V D S P D G F V A H P D L
15 TCGACGCGGTCTTCTCCGCGGTCCGGCAGGGAGCCGCCAGCCGACCGGA 3500
L D A V F S A V G D G S R Q P T G
TGGCGCGACCTCGCGGTGCACGCGTCCGACGCCACCGTGTGCGCGCCTG 3550
W R D L A V H A S D A T V L R A C
CCTCACCCGCGCGACAGTGGTGTCTGAGCTCGCCGCCTTCGACGGTG 3600
20 L T R R D S G V V E L A A F D G
CCGGAATGCGGTGCTCACCGCGGAGTCGGTGACGCTGGGCGAGGTCCGG 3650
A G M P V L T A E S V T L G E V A
TCGGCAGGCGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGTT 3700
S A G G S D E S D G L L R L E W L
25 GCCGGTGGCGGAGGGCCACTACGACGGTGCCGACGAGTCCCGAGGGCT 3750
P V A E A H Y D G A D E L P E G
ACACCCTCATCACCGCCACACACCCCGACGACCCCGACGACCCCAAC 3800
Y T L I T A T H P D D P D P T N
CCCCACAACACACCCACACGCACCCACACACAAACCACACGCGTCTCAC 3850
30 P H N T P T R T H T Q T T R V L T
CGCCCTCCAACACCACCTCATCACCACCAACCACACCCTCATCGTCCACA 3900
A L Q H H L I T T N H T L I V H
CCACCACCGACCCCCCAGGCGCCGCGGTACCGGCCTCACCCGACCCGCA 3950
T T T D P P G A A V T G L T R T A
35 CAAAACGAACACCCCGCGCATCCACCTCATCGAAACCCACACCCCA 4000
Q N E H P G R I H L I E T H H P H
CACCCCACTCCCCCTCACCAACTCACACCCTCCACCAACCCCACTAC 4050
T P L P L T Q L T T L H Q P H L
GCCTCACCAACAACACCTCCACACCCCCACCTCACCCCATCACCAAC 4100
40 R L T N N T L H T P H L T P I T T
CACCACAACACCACCAACCACCCCAACACCCCAACCCCTCAACCCCA 4150
H H N T T T T P N T P P L N P N
CCACGCCATCCTCATCACCGGCGGCTCCGGCACCCCTCGCCGGCATCCTCG 4200
H A I L I T G G S G T L A G I L
45 CCGGCCACCTCAACCACCCCAACACCTACCTCCTCTCCCGCACACCACCA 4250
A R H L N H P H T Y L L S R T P P
CCCCCACCACACCCGGCACCCACATCCCTGCGACCTCACCGACCCCA 4300
P P T T P T H I P C D L T D P T
CCAAATCACCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCT 4350
50 Q I T Q A L T H I P Q P L T G I
TCCACACCGCGCCACCCCTCGACGACGCCACCCCTCACCAACCTCACCC 4400
F H T A A T L D D A T L T N L T P
CAACACCTCACCAACCCCTCAACCCAAAGCCGACGCCGCTGGCACCT 4450
Q H L T T T L Q P K A D A A W H L
55 CCACCACCAACCCAAACCAACCCCTCACCACTTCGTCTCTACTCCA 4500
H H H T Q N Q P L T H F V L Y S
GCGCGCGCGCCACCCCTCGGCAGCCCCGGCCAAGCCAACCTACGCCGCGCC 4550
S A A A T L G S P G Q A N Y A A A
AACGCCTTCCTCGACGCCCTCGCCACCCACCGCCACACCAAGGACAAC 4600
60 N A F L D A L A T H R H T Q G Q P

CGCCACCACCATCGCCTGGGGCATGTGGCACACCACCACACTCACCA 4650
 A T T I A W G M W H T T T T L T
 GCCAACTCACCGACAGCGACCGCGACCGCATCCGCGCGGGCGGCTTCCTG 4700
 S Q L T D S D R D R I R R G G F L
 5 CCGATCTCGGACGACGAGGGCATGC
 P I S D D E S M

The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

10 GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
 M R L Y E A A R R T G S P V V V
 GCGGCGCGGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
 A A A L D D A P D V P L L R G L R
 GCGTACGACCGTCCGGCGTGCCGCGGTCCGGGAACGCTCTCTCGCCGACC 150
 15 R T T V R R A A V R E R S L A D
 GCTCGCGGTGCTGCCGACGACGAGCGCGCGGACGCCTCCCTCGCGTTTCG 200
 R S P C C P T T S A P T P P S R S
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
 S W N S T A T V L G H L G A E D I
 20 CCGGCGGACGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300
 P A T T T F K E L G I D S L T A
 TCCAGCTGCGCAACGCGCTGACCAACGGCGACCGGCGTACGCCTCAACGCC 350
 V Q L R N A L T T A T G V R L N A
 ACAGCGGTCTTCGACTTTCGACCGCGCGCGGCTCGCGCGAGACTCGG 400
 25 T A V F D F P T P R A L A A R L G
 CGACGAGCTGGCCGGTACCCGCGCGCCCGTTCGCGGCCCGGACCGCGGCCA 450
 D E L A G T R A P V A A R T A A
 CCGCGGCGCGGACGACGACGACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
 T A A A H D E F L A I V G M A C R
 30 CTGCGCGGGCGGGTTCGCGTCCGACAGGAGCTGTGGCGTCTCGTCGCGTC 550
 L P G G V A S P Q E L W R L V A S
 CGGCACCGACGCCATCACGGAGTTCGCCGCGGACCGCGGCTGGGACGTGG 600
 G T D A I T E F P A D R G W D V
 ACGCGCTCTACGACCCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
 35 D A L Y D P D P D A I G K T F V R
 CACGGCGGCTTCCTCGACGGTGCAGCGGCTTCGACGCGGCGTTCTTCGG 700
 H G G F L D G A T G F D A A F F G
 GATCAGCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
 I S P R E A L A M D P Q Q R V L
 40 TGGAGACGTCTCGGAGGCGTTTCGAAAGCGCGGGCATCACCCCGGACGCG 800
 L E T S W E A F E S A G I T P D A
 GCGCGGGGACGACACCGCGGTTCATCGGCGCGTTCTCCTACGGGTA 850
 A R G S D T G V F I G A F S Y G Y
 CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGACCA 900
 45 G T G A D T N G F G A T G S Q T
 GCGTGCTCTCGGCGCGCCTCTCTACTTCTACGGTCTGGAGGGCCCTTCG 950
 S V L S G R L S Y F Y G L E G P S
 GTCACGGTCGACACCGCCTGCTCGTCACTGGTCCGCTGCACCAGGC 1000
 V T V D T A C S S S L V A L H Q A
 50 AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTCCGCGGTG 1050
 G Q S L R S G E C S L A L V G G
 TCACGGTGATGGCGTCCGCGCGGCGGATTCGTGAGTTCTCCCGGACGCGC 1100
 V T V M A S P G G F V E F S R Q R
 GGGCTCGCGCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
 55 G L A P D G R A K A F G A G A D G
 TACGAGCTTCGCGAGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
 T S F A E G A G A L V V E R L S
 ACGCGGAGCGCCACGGCCACCGCTCCTCGCCCTCGTACGCGGCTCCGCG 1250
 D A E R H G H T V L A L V R G S A

GCTAACTCCGACGGCGCGTCGAACGGTCTGTGCGGCGCGGAACGGCCCCCTC 1300
A N S D G A S N G L S A P N G P S
CCAGGAACCGGTCCATCCACCAGGCCCTCGCGAACCGGAACTCACCCTCG 1350
Q E R V I H Q A L A N A K L T P
5 CCGATGTCCAGCGCGTCCGAGGCGCACGGCACCGGACCCGCGCTCGGCGAC 1400
A D V D A V E A H G T G T R L G D
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
P I E A Q A L L A T Y G Q D R A T
10 GCCCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCCAGGCCG 1500
P L L L G S L K S N I G H A Q A
CGTCAGGGGTGCGCGGATCATCAAGATGGTGACGGCCATCCGGCACGGG 1550
A S G V A G I I K M V Q A I R H G
GAACTGCCGCGGACACTGCACGCGGACGCCCTCGCGCACGCTCGACTG 1600
E L P P T L H A D E P S P H V D W
15 GACGGCCGGTGCCTGAGCTCCTGACGTGCGGCCGCGCGTGGCCGGGGA 1650
T A G A V E L L T S A R P W P G
CCGCTCGCCCTAGGCGGGCAGGCGTGTGCTCCTTCGGGATCAGTGGCACC 1700
T G R P R R A G V S S F G I S G T
AAGCCCCAGTCCCTGGAAGCGCACCCCCCACTCAGCCTGCGGACAA 1750
20 N A H V I L E S A P P T Q P A D N
CGCGGTGATCGAGCGGGCACCAGGAGTGGGTGCGGTTGGTGATTTCGGCCA 1800
A V I E R A P E W V P L V I S A
GGACCCAGTCCGCTTTGACTGAGCACGAGGGCCGGTTGCGTGCGTATCTG 1850
R T Q S A L T E H E G R L R A Y L
25 CCGGCGTCCCGGGGTGGATATGCGGGCTGTGGCATCGACGCTGGCGAT 1900
A A S P G V D M R A V A S T L A M
GACACGGTCCGCTGTTGAGCACCGTGCCGTGCTGCTGGGAGATGACACCG 1950
T R S V F E H R A V L L G D D T
30 TCACCGGCACCGCTGTGTCTGACCCTCGGGCGGTGTTGCTCTTCCCGGGA 2000
V T G T A V S D P R A V F V F P G
CAGGGGTCCGACGCTGCTGGCATGGGTGAGGAACTGGCCGCCGCGTTCCC 2050
Q G S Q R A G M G E E L A A A F P
CGTCTTCGCGCGATCCATCAGCAGGTGTGGGACCTGCTCGATGTGCCCCG 2100
V F A R I H Q Q V W D L L D V P
35 ATCTGGAGGTGAACGAGACCGGTTACGCCAGCCGGCCCTGTTTCGCAATG 2150
D L E V N E T G Y A Q P A L F A M
CAGGTGGCTCTGTTCCGGGCTGCTGGAATCGTGGGGTGTACGACCGGACGC 2200
Q V A L F G L L E S W G V R P D A
GGTGATCGGCCATTCGGTGGGTGAGCTTGCGGCTGCGTATGTGTCCGGGG 2250
40 V I G H S V G E L A A Y V S G
TGTGGTCTGTTGGAGGATGCTGCACCTTGGTGTGCGGCGGGCTCGTCTG 2300
V W S L E D A C T L V S A R A R L
ATGCAGGCTCTGCCCGCGGGTGGGGTGATGGTCTGCTGCCCGGTCTCGGA 2350
M Q A L P A G G V M V A V P V S E
45 GGATGAGGCCCGGGCCGTGCTGGGTGAGGGTGTGGAGATCGCCGCGGTCA 2400
D E A R A V L G E G V E I A A V
ACGGCCCGTCCGTGGTGTCTCTCCGGTGATGAGGCCCGCGTGTGCTGACG 2450
N G P S S V V L S G D E A A V L Q
GCCCGGAGGGGTGGGGAAGTGACGCGGCTGGCGACCGACCGCGTT 2500
50 A A E G L G K W T R L A T S H A F
CCATTCGCCCGTATGGAACCCATGCTGGAGGAGTTCCGGGCGGTGCGCG 2550
H S A R M E P M L E E F R A V A
AAGGCCTGACCTACCGGACCGCGCAGGTCTCCATGGCCGTTGGTGATCAG 2600
E G L T Y R T P Q V S M A V G D Q
55 GTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGGACACGGTCCGGTT 2650
V T T A E Y W V R Q V R D T V R F
CGGCGAGCAGGTGGCCTCGTACGAGGACGCCGTGTTGCTCGAGCTGGGTG 2700
G E Q V A S Y E D A V F V E L G
CCGACCGGTCACTGGCCCGCTGGTTCGACGGTGTGCGGATGCTGCACGGC 2750
60 A D R S L A R L V D G V A M L H G

SUBSTITUTE SHEET (RULE 26)

GACCACGAAATCCAGGCCGCGATCGGCGCCCTGCCCCACCTGTATGTCAA 2800
 D H E I Q A A I G A L A H L Y V N
 CGGGGTACGGTCCGACTGGCCCCGCGCTCCTGGGCGATGCTCCGGCAACAC 2850
 S T T V D W P A L L G D A P A T
 5 GGGTGTCTGACCTTCCGACATACGCCTTCCAGCACCAGCGCTACTGGCTC 2900
 R V L D L P T Y A F Q H Q R Y W L
 GAGTCGGCTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCTCGGCAC 2950
 E S A P P A T A D S G H P V L G T
 CGGASTCGCGCTCCCCGGGTGCGCGGGCCGGGTGTTACGGGTCCCGTGC 3000
 10 S Y A V A G S P G R V F T G P V
 CCGCCGCTGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCC 3050
 P A G A C R A V F I A E L A L A A
 GCGGACGCCACCGACTCGGCCACGGTCTGAACAGCTCGACGTCACCTCCGT 3100
 A D A T D C A T V E Q L D V T S V
 15 GCGCGGCGGATCCGCCCCGCGGCAGGGCCACCGCGCAGACCTGGGTGATG 3150
 F G G S A R G R A T A Q T W V D
 AACCCCGCGCGGACGGGCGGCGCGCTTACCGTCCACACCCGCGTCCGC 3200
 E P A A D G R R R F T V H T R V G
 GACGCCCCGTGGACGCTGCACGCCGAGGGGGTTCTCCGCCCCGCGCGCT 3250
 20 D A P W T L H A E G V L R P G R V
 GCGCCAGCCCGAAGCCGTGACACCGCTGGCCCCCGCGGGCGCGGTGC 3300
 F Q P E A V D T A W P P P G A V
 CCGCGGACGGGCTGCCCCGGGCGTGGCGACGCGCGGACCAGGTCTTCGTC 3350
 P A D G L P G A W R R A D Q V F V
 25 GAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCT 3400
 E A E V D S P D G F V A H P D L L
 CGACGCGGTCTTCTCCGCGGTGCGCGACGGGAGCCGCCAGCCGACCGGAT 3450
 D A V F S A V G D G S R Q P T G
 GGCGGACCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCCTGC 3500
 30 W R D L A V H A S D A T V L R A C
 CTCACCCCGCGGACAGTGGTGTGCTGGAGCTCGCCGCTTCGACGGTGC 3550
 L T R R D S G V V E L A A F D G A
 CGGAATGCCGGTGTCTACCGCGGAGTGGTGACGCTGGGCGAGGTCCGCT 3600
 G M P V L T A E S V T L G E V A
 35 CGGCAGGCGGATCCGACGAGTCGGACGGTCTGCTTCGGCTTGAGTGGTTG 3650
 S A G G S D E S D G L L R L E W L
 CCGGTGGCGGAGGCCCCACTACGACGGTGCCGACGAGCTGCCCCGAGGGCTA 3700
 P V A E A H Y D G A D E L P E G Y
 CACCCCTCATCAGCCACACACCCCGACGACCCCGACGACCCCAACCAACC 3750
 40 T L I T A T H P D D P D D P T N
 CCCACAACACACCCACACGCACCCACACACAAACCACACGGTCTCTCACC 3800
 P H N T P T R T H T Q T T R V L T
 GCCCTCCAACACCACCTCATCACCACCAACCACACCCCTCATCGTCCACAC 3850
 A L Q H H L I T T N H T L I V H T
 45 CACCACGACCCCCCAGGCGCGCGCTCACCAGCCTCACCAGCACCGCAC 3900
 T T D P P G A A V T G L T R T A
 AAAACGAACACCCCGCGCGCATCCACCTCATCGAAACCCACACCCCCAC 3950
 Q N E H P G R I H L I E T H H P H
 ACCCCACTCCCCCTCACCCTCACTCACCACCTCCACCAACCCCACTACG 4000
 50 T P L P L T Q L T T L H Q P H L R
 CCTCAGCAACAACCCCTCCACACCCCCACCTCACCCTCATCACCACCC 4050
 E T N N T L H T P H L T P I T T
 ACCACAACACCAACCAACCAACCCCAACACCCCAACCCCTCAACCCCAAC 4100
 H H N T T T T T P N T P P L N P N
 55 CACGCGATCCTCATCACCAGGCGGCTCCGGCACCCCTCGCGGCGATCCTCGC 4150
 H A I L I T G G S G T L A G I L A
 CCGGACCTCAACACCCCAACACCTACCTCCTCTCCCGCACACCAAC 4200
 R H L N H P H T Y L L S R T P P
 CCGGACCAACCCCGGACCCACATCCCTGCGACCTCACCAGCCCAAC 4250
 60 P P T T P G T H I P C D L T D P T

SUBSTITUTE SHEET (RULE 26)

CAAATCACCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCTT 4300
 Q I T Q A L T H I P Q P L T G I F
 CCACACCGCGCGCCACCCTCGACGACGCCACCCTCACCAACCTCACCCCCC 4350
 H T A A T L D D A T L T N L T P
 5 AACACCTCACCAACCCTCCAACCCAAAGCCGACGCCGCTGGCACCTC 4400
 Q H L T T T L Q P K A D A A W H L
 CACCACCACACCCAAAACCAACCCCTCACCACTTCGTCTCTACTCCAG 4450
 H H H T Q N Q P L T H F V L Y S S
 CGCGCGCGCCACCCTCGGCAGCCCCGGCCAAGCCAACTACGCCGCGCCA 4500
 10 A A A T L G S P G Q A N Y A A A
 ACGCCTTCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCC 4550
 N A F L D A L A T H R H T Q G Q P
 GCCACCACCTCGCCTGGGGCATGTGGCACACCACCACCACTCACCAG 4600
 A T T I A W G M W H T T T T L T S
 15 CCAACTACCGACAGCGACCGCGACCGCATCCGCGCGGGGCTTCCTGC 4650
 Q L T D S D R D R I R R G G F L
 CGATCTCGGACGACGAGGGCATGC
 P I S D D E G M

20 The *AvrII-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module
 13 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
 M R L Y E A A R R T G S P V V V
 GCGGCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
 25 A A A L D D A P D V P L L R G L R
 GCGTACGACCGTCCGGCGTGCCGCGCGTCCGGGAACGCTCTCTCGCCGACC 150
 R T T V R R A A V R E R S L A D
 GCTCGCGGTGCTGCCCGACGACGAGCGCGCGGACGCTCCCTCGCGTTCTG 200
 R S P C C P T T S A P T P P S R S
 30 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCGGAAGACAT 250
 S W N S T A T V L G H L G A E D I
 CCGGCGGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300
 P A T T T F K E L G I D S L T A
 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGGCTACGCCTCAACGCC 350
 35 V Q L R N A L T T A T G V R L N A
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCTCGCCGCGAGACTCGG 400
 T A V F T P R A L A R L G
 CGACGAGCTGCCCGGTACCCGCGCGCGCGTTCGCGGCGCGGACCGCGGCCA 450
 D E L A G T R A P V A A R T A A
 40 CCGCGCGCGCGCACGACGAACCGTGCGGATCGTGCGCATGGCCTGCCGT 500
 T A A A H D E P L A I V G M A C R
 CTGCCGGGCGGGGTCCGCTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
 L P G G V A S P Q E L W R L V A S
 CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
 45 G T D A I T E F P A D R G W D V
 ACGCGCTCTACGACCCGGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
 D A L Y D P D P D A I G K T F V R
 CACGGCGGCTTCCTCGACCGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
 H G G F L D G A T G F D A A F F G
 50 GATCAGCCCGCGGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
 I S P R E A L A M D P Q Q R V L
 TGGAGACGTCTGGGAGGCGTTGGAAGCGCGGGCATACCCCGGACGCG 800
 L E T S W E A F E S A G I T P D A
 GCGCGGGGCGAGACACCGCGGTGTTGATCGGCGCGTTCTCTACGGGTA 850
 55 A R G S D T G V F I G A F S Y G Y
 CGGCACGGGTGCGGATACCAAGCGCTTCGGCGGACAGGGTTCGACACCA 900
 G T G A D T N G F G A T G S Q T
 GCGTGCTCTCGGCGCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
 S V L S G R L S Y F Y G L E G P S

CTCACGGTTCGACACCGCCTGCTCGTCTCGTCACTGGTCCGCCCTGCACCAGGC 1000
V T V D T A C S S S L V A L H Q A
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050
G Q S L R S G E C S L A L V G G
5 TCACGGTGATGGCGTCGCCCCGGGATTTCGTTCGAGTTCTCCCGGCAGCGC 1100
V T V M A S P G G F V E F S R Q R
GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
G L A P D G R A K A F G A G A D G
TACGAGCTTCGCCGAGGGCGCCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200
10 T S F A E G A G A L V V E R L S
ACGCGGAGCGCCACGGCCACACCGTCCTCGCCCTCGTACGCGGCTCCGCG 1250
D A E R H G H T V L A L V R G S A
GCTAACTCCGACGGCGCGTCGAACGGTCTGTTCGGCGCCGAACGGCCCCCTC 1300
A N S D G A S N G L S A P N G P S
15 CCAGGAACGCGTCATCCACCAGGCCCTCGCGAACCGGAAACTCACCCCCG 1350
Q E R V I H Q A L A N A K L T P
CCGATGTTCGACGCGGTTCGAGGCGCACGGCACCGGCCCCCGCTCGGCGAC 1400
A D V D A V E A H G T G T R L G D
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
20 P I E A Q A L L A T Y G Q D R A T
GCCCTGCTGCTCGGCTCGCTGAAGTCGAACATCGGGCACGCCAGGCCG 1500
P L L L G S L K S N I G H A Q A
CGTCAGGGGTTCGCCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
A S G V A G I I K M V Q A I R H G
25 GAACGTCCGCGGACACTGCACGCGGACGAGCCGTTCGCCGACGTCGACTG 1600
E L P P T L H A D E P S P H V D W
GACGGCCGGTTCGCGTCGAGCTCCTGACGTTCGGCCCGGCCGTGGCCGGGGA 1650
T A G A V E L L T S A R P W P G
CCGGTTCGCCCTAGGCGGGCGGGCGTGTCTGCTCCTTCGGAGTCAGCGGCACC 1700
30 T G R P R R A G V S S F G V S G T
AACGCCCCACGTATCCTGGAGAGCGCACCCCCCGCTCAGCCCCGCGGAGGA 1750
N A H V I L E S A P P A Q P A E E
GGCGCAGCCTGTTGAGACGCCGGTGGTGGCCTCGGATGTGCTGCCGCTGG 1800
A Q P V E T P V V A S D V L P L
35 TGATATCGGCCAAGACCCAGCCCGCCCTGACCGAACACGAAGACCGGCTG 1850
V I S A K T Q P A L T E H E D R L
CGCGCCTACCTGGCGGCGTTCGCCCGGGGCGGATATACGGGCTGTGGCATC 1900
R A Y L A A S P G A D I R A V A S
36 GACCGTGGCGGTGACACGGTTCGTTTCGAGCACCGCGCGTACTCCTTG 1950
40 T L A V T R S V F E H R A V L L
GAGATGACACCGTCAACGGCACCGCGGTGACCGACCCCAGGATCGTGTTC 2000
G D D T V T G T A V T D P R I V F
GTCTTTCCCGGGCAGGGGTGGCAGTGGCTGGGGATGGGCAGTGCAGTGGC 2050
V F P G Q G W Q W L G M G S A L R
45 CGATTTCGTCGGTGGTTCGCGGAGCGGATGGCCGAGTGTGCGGCGGCGT 2100
D S S V V F A E R M A E C A A A
TGCGCGAGTTTCGTGGACTGGGATCTGTTCACGGTTCTGGATGATCCGGCG 2150
L R E F V D W D L F T V L D D P A
GTGGTGGACCGGGTTGATGTGGTCCAGCCCGCTTCCTGGGCGATGATGGT 2200
50 V V D R V D V V Q P A S W A M M V
TTCCCTGGCCCGGGTGTGGCAGGCGGCCGGTGTGCGGCCGGATGCGGTGA 2250
S L A A V W Q A A G V R P D A V
TCGGCCATTTCGAGGGTGAGATCGCCCGCAGCTTGTGTGGCGGGTSCGGTG 2300
I G H S Q G E I A A A C V A G A V
55 TCACTACCGGATGCGGCCCGGATCGTGACCTTGCGCAGCCAGGCGATCGC 2350
S L R D A A R I V T L R S Q A I A
CCGSGGCCTTCGCGGCCCGGGGCGGATGGCATCCGTCCGCCCTGCCCGGCC 2400
E G L A G R G A M A S V A L P A
AGGATGTTCGAGCTGGTTCGACGGGGCCTGGATCGCCGCCACAAACGGGGCC 2450
60 L D V E L V D G A W I A A H N G P

SUBSTITUTE SHEET (RULE 26)

GCCTCCACCGGTGATCGCGGGCACCOCGGAAGCGGTGACCATGTCTCAC 2500
A S T V I A G T P E A V D H V L T
CGCTCATGAGGCACAAGGGGTGCGGGTGGCGGGATCACCGTGGACTATG 2550
A H E A G V R V R R I T V D Y
5 CCTCGCACACCCCGCACGTGAGCTGATCCGCGACGAACACTACTCGACATC 2600
A S H T P H V E L I R D E L L D I
ACTAGCGACAGCAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTGACCGT 2650
T S D S S S Q T P L V P W L S T V
GGACGGCACCTGGTTCGACAGCCCGCTGGACGGGGAGTACTGGTACCGGA 2700
10 D G T W V D S P L D G E Y W Y R
ACCTGCGTGAACCGGTGCGTTTCCACCCCGCGCTCAGCCAGTTGCAGGCC 2750
N L R E P V G F H P A V S Q L Q A
CAGGCGCACCGTGTTCGTGAGGTGAGCGCCAGCCCGGTGTTGTTGCA 2800
Q G D T V F V E V S A S P V L L Q
15 GGCGATGGACGACGATGTCGTCACGGTTGCCACGCTGCGTCGTGACGACG 2850
A M D D D V V T V A T L R R D D
GCGACGCTACCCGATGCTCACC GCCCTGGCACAGGCCTATGTCCACGGC 2900
G D A T R M L T A L A Q A Y V H G
GTCACCGTGGACTGGCCCGCATCCTCGGCACCAACACCCGGGTACT 2950
20 V T V D W P A I L G T T T T R V L
GGACCTTCCGACCTACGCCCTTCCAACACAGCGGTACTGGCTCGAGTCGG 3000
D L P T Y A F Q H Q R Y W L E S
CTCCCCCGGCCACGGCCGACTCGGGCCACCCCGTCTCGGCACCGGAGTC 3050
A P P A T A D S G H P V L G T G V
25 GCGGTGCGCGGGTTCGCGGGCGGGTGTTCACGGGTCCCGTGCCCGCCGG 3100
A V A G S P G R V F T G P V P A G
TGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCCGCCGACG 3150
A D R A V F I A E L A L A A A D
CCACCGACTGCGCCACGGTCAACAGCTCGACGTCACCTCCGTGCCCGGC 3200
30 A T D C A T V E Q L D V T S V P G
GGATCCGCCCCGCGCAGGGCCACCGCGCAGACCTGGGTGATGAACCCGC 3250
G S A R G R A T A Q T W V D E P A
CGCCGACGGGGCGCGCGCTTACCGTCCACACCCCGCTCGGCGACGCC 3300
A D G R R R F T V H T R V G D A
35 CGTGGAGCTGCACGCGGAGGGGTCTCCGCCCCGCGCGTGGCCCGAG 3350
P W T L H A E G V L R P G R V P Q
CCCGAAGCGGTGACACCGCTGGCCCCCGCGGGCGCGGTGCCCCGCGGA 3400
P E A V D T A W P P P G A V P A D
CGGGGTGCCCCGGGGCGTGGCGACGCGCGGACAGGTCTTCGTGGAAGCCG 3450
40 G L P G A W R R A D Q V F V E A
AAGTCGACAGCCCTGACGGCTTCGTGGCACACCCCGACCTGCTCGACGCG 3500
E V D S P D G F V A H P D L L D A
GTCTTCTCCGCGGTGCGGACGCGGAGCCGCCAGCCGACCGGATGGCGCGA 3550
V F S A V G D G S R Q P T G W R D
45 CCTCGCGGTGCACGCGTGGACGCCACCGTGTGCGCGCCTGCCTCACCC 3600
L A V H A S D A T V L R A C L T
GCCGCGACAGTGGTGTGCTGGAGCTCGCCGCTTCGACGGTGCCGGAATG 3650
R R D S G V V E L A A F D G A G M
CCGGTGTCTACCGCGGAGTGGTGACGCTGGGCGAGGTGCGGTGCGGCGAG 3700
50 P V L T A E S V T L G E V A S A G
CGGATCCGACGAGTGGACGGTCTGCTTCGGCTTGAGTGGTGGCCGGTGG 3750
G S D E S D G L L R L E W L P V
CGGAGGCCCACTACGACGGTGCCGACGAGCTGCCCGAGGGCTACACCCCTC 3800
A E A H Y D G A D E L P E G Y T L
55 ATCACCGCCACACACCCCGACGACCCCGACGACCCCAACCCCAACAA 3850
I T A T H P D D P D D P T N P H N
CACACCCACACGACCCCAACAAACACACGCGTCTCACCCCTCTCC 3900
T P T R T H T Q T T R V L T A L
AACACCACCTCATCACCAACCAACACACCTCATCGTCCACACCAACC 3950
60 Q H H L I T T N H T L I V H T T T

SUBSTITUTE SHEET (RULE 26)

GACCCCCCAGGGCGCGCGCTCACCGGCCTCACCCGCACCGCACAAAACGA 4000
 D P P G A A V T G L T R T A Q N E
 ACACCCCGGGCATCCACCTCATCGAAACCCACCCACCCACACCCAC 4050
 H P G R I H L I E T H E P H T F
 5 TCCCTCTACCCCACTCACCACCTCCACCAACCCACCTACGCTCACC 4100
 L P L T Q L T T L H Q P H L R L T
 AACAAACACCTCCACACCCCGCACCTCACCCCATCACCACCCACCAAA 4150
 N N T L H T P H L T P I T T H H N
 CACCACCAACCCACCCCAACACCCACCCCTCAACCCCAACCCACGCCA 4200
 10 T T T T T P N T P P L N P N H A
 TCCTCATCACCGCGCGCTCCGGCACCTCGCGGCATCCTCGCTCGCCAC 4250
 I L I T G G S G T L A G I L A R H
 CTCACCCACCCACACCTACCTCCTCTCCCGCACACCACCCACCCAC 4300
 L N H P H T Y L L S R T P P P P T
 15 CACACCCGGCACCCACATCCCTGCGACCTCACCGACCCACCCAAATCA 4350
 T P G T H I P C D L T D P T Q I
 CCCAAGCCCTCACCCACATACCACAACCCCTCACCGGCATCTTCCACACC 4400
 T Q A L T H I P Q P L T G I F H T
 GCGCCACCCCTCGACGACGCCACCCCTCACCAACCTCACCCCCCAACACCT 4450
 20 A A T L D D A T L T N L T P Q H L
 CACCACCCCTCCAACCCAAAGCCGACGCCGCTGGCACCTCCACCACC 4500
 T T T L Q P K A D A A W H L H H
 ACACCCAAAACCAACCCCTCACCCACTTCGTCTCTACTCCAGCGCCGCC 4550
 H T Q N Q P L T H F V L Y S S A A
 25 GCCACCCCTCGGCAGCCCGGCCAAGCCAAGCTACGCCGCGCAACGCCTT 4600
 A T L G S P G Q A N Y A A A N A F
 CCTCGACGCCCTCGCCACCCACCGCCACACCCAAGGACAACCCGCCACCA 4600
 L D A L A T H R H T Q G Q P A T
 CCATCGCCTGGGGCATGTGGCACACCACCACACTCACCAGCCAACTC 4700
 30 T I A W G M W H T T T T L T S Q L
 ACCGACAGCGACCGGACCGCATCCGCCGCGCGGCTTCCTGCGGATCTC 4750
 T D S D R I R R G G F L P I S
 GGACGACGAGGGCATGC
 35 D D E G M

The *NheI-XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module 12 of rapamycin is shown below.

GCATGCGGCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTGGTG 50
 M R L Y E A A R R T G S P V V V
 40 CGGCGCGGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGGCTGCG 100
 A A A L D A P D V P L L R G L R
 GCGTACGACCGTCCGGCGTGCCGCGCTCCGGGAACGCTCTCTCGCCGACC 150
 R T T V R R A A V R E R S L A D
 45 GCTCGCGGCTGCTGCCCCGACGACGAGCGCGCGGACGCTCCCTCGCGTTCC 200
 R S P C C P T T S A P T P P S R S
 TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
 S W N S T A T V L G H L G A E D I
 CCGGCGACGACGACGTTCAAGGAACTCGGCATCGACTCGCTCACCGCGG 300
 P A T T T F K E L G I D S L T A
 50 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGCC 350
 V Q L R N A L T T A T G V R L N A
 ACAGCGGTCTTCGACTTTCCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400
 T A V F D F P T P R A L A A R L G
 CGACGAGCTGGCCGTTACCCGCGCGCCCGTCCGCGCCCGGACCGCGGCCA 450
 55 D E L A G T R A P V A A R T A A
 CCGCGCGCGCGCACGACGACCGCTGGCGATCGTGGGATGGGCTCCCGT 500
 T A A A H D E P L A I V G M A C R
 CTGCCGCGCGGGGTGCGTCCGACAGGAGCTGTGGCGTCTCGTCCGCTC 550
 L P G G V A S P Q E L W R L V A S

CGGCACCGACGCCATCACGGAGTTCCCCGCGGACCGCGGCTGGGACGTGG 600
G T D A I T E F P A D R G W D V
ACGCGCTCTACGACCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
D A L Y D P D F D A I G K T F V R
5 CACGGCGGCTTCCTCGACGGTSCGACCGGCTTCGACCGCGGCTTCCTCGG 700
H G G F L D G A T G F D A A F F G
GATCAGCCCCGCGAGGCCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
I S P R E A L A M D P Q Q R V L
TGGAGACGTCTGGGAGGCGTTGGAAGCGCGGGCATCACCCCGGACGCG 800
10 L E T S W E A F E S A G I T P D A
GCGCGGGGACGACACCGGCGTGTTCATCGGCGCGTTCCTACGGGTGTA 850
A R G S D T G V F I G A F S Y G Y
CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGACAGCCA 900
G T G A D T N G F G A T G S Q T
15 GCGTGCTCTCCGGCCGCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
S V L S G R L S Y F Y G L E G P S
GTCACGGTGCACACCGCCTGCTCGTCTGCTACTGGTGGCCCTGCACAGGC 1000
V T V D T A C S S S L V A L H Q A
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050
20 G Q S L R S G E C S L A L V G G
TCACGGTGATGGCGTCCGCGGCGATTGCTCGAGTTCTCCCGGACGCGC 1100
V T V M A S P G G F V E F S R Q R
GGGCTCGCGCCGGACGGGCGGGCGAAGGCGTTCGGCGCGGGCGCGGACGG 1150
G L A R D G R A K A F G A G A D G
25 TACGAGCTTCGCGGAGGGGCGCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200
T S F A E G A G A L V V E R L S
ACGCGGAGCGCCACGGCCACACCGTCTCTCGCCCTCGTACGCGGCTCCGCG 1250
D A E R H G H T V L A L V R G S A
GCTAACTCCGACGGCGCGTGAACGGTCTGTGCGCGCGGAACGGCCCTC 1300
30 A N S D G A S N G L S A P N G P S
CCAGGAACGCGTATCCACAGGCCCTCGCGAACGCGAACTCACCCCCG 1350
Q E R V I H Q A L A N A K L T P
CCGATGTGACGCGGTTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400
A D V D A V E A H G T G T R L G D
35 CCCATCGAGGCGCAGGCGTGTCTCGCGACGTACGGACAGGACCGGGCGAC 1450
P I E A Q A L L A T Y G Q D R A T
GCCCCGTGCTGCTCGGCTCGCTGAAGTGAACATCGGGCACGCCAGGCCG 1500
P L L L G S L K S N I G H A Q A
CGTCAGGGGTGCGCGGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
40 A S G V A G I I K M V Q A I R H G
GAACTGCCCGGACACTGCACGCGGACGCGTTCGCCGACGTCGACTG 1600
E L P P T L H A D E P S P H V D W
GACGGCCGGTGGCGTTCGAGCTCCTGACGTGCGCCCGGCGGTGGCCGGGA 1650
T A G A V E L L T S A R P W P G
45 CCGGTGCGCCGCGCGCGCTGCCGTCTCGTCTGTTGGCGTGAGCGGCACG 1700
T G R P R R A A V S S F G V S G T
AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTGCA 1750
N A H I I L E A G P V K T G P V E
GGCAGGAGCGATCGAGGCAGGACCGGTGAAGTAGGACCGGTTCGAGGCTG 1800
50 A G A I E A G P V E V G P V E A
GACCGCTCCCGCGGCGCGCGCTCAGCACCGGGCGAAGACCTTCGGCTG 1850
G P L P A A P P S A P G E D L P L
CTCGTGTGCGCGCGTTCGCCGAGGCACTCGACGAGCAGATCGGGCGCCT 1900
L V S A R S P E A L D E Q I G R L
55 GCGCGCCTATCTCGACACCGGCGCGGCGTTCGACCGGGCGCGCGTGGCGC 1950
R A Y L D T G P G V D R A A V A
AGACACTGGCCCGGCGTACGCACTTCACCCACCGGGCGTACTGCTCGGG 2000
Q T L A R R T H F T H R A V L L G
GACACCGTTCATCGGCGCTCCCCCGCGGACAGGCGGACGAAGTCTCTT 2050
60 D T V I G A P P A D Q A D E L V F

SUBSTITUTE SHEET (RULE 26)

CGTCTACTCCGGTCAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100
V Y S G Q G T Q H P A M G E Q L
CCGCCCGGTTCCCCGTCTTCGCGCGGATCCATCAGCAGGTGTGGGACCTG 2150
A A A F P V F A R I H Q Q V W D L
5 CTGGATGTGCCCGATCTGAGGTGAACGAGACCGGTTACGCCCAGCCGGC 2200
L D V P D L E V N E T G Y A Q P A
CCTGTTCCGAATGCAGCTGCTCTGTTCCGGGCTGCTGGAATCGTGGGGTG 2250
L F A M Q V A L F G L L E S W G
TACGACCGGACCGGTTGATCGGCCATTCGGTGGGTGAGCTTGCGGCTGCG 2300
10 T R P D A V I G H S V G E L A A A
TATGTGTCCGGGGTGTGGTCTGGAGGATGCCTGCACTTTGGTGTCCGGC 2350
Y V S G V W S L E D A C T L V S A
GCGGGCTCGTCTGATGCAGGCTCTGCCCCGCGGGTGGGGTGATGGTCTGCTG 2400
R A R L M Q A L P A G G V M V A
15 TCCCGGTCTCGGAGGATGAGGCCCGGGCCGTGCTGGGTGAGGGTGTGGAG 2450
V P V S E D E A R A V L G E G V E
ATCGCCGCGGTCAACGGCCCGTCTGCTGGTGTCTCTCCGGTGTGAGGC 2500
I A A V N G P S S V V L S G D E A
CGCCGTGCTGACGGCCCGGAGGGGCTGGGGAAGTGGACGCGGCTGGCGA 2550
20 A V L Q A A E G L G K W T R L A
CCAGCCACGCGTTCCATTCCGCCCGTATGGAACCCATGCTGGAGGAGTTC 2600
T S H A F H S A R M E P M L E E F
CGGGCGGTGCGCCGAAGGCCTGACCTACCGGACGCGCAGGTCTCCATGGC 2650
R A V A E G L T Y R T P Q V S M A
25 CGTTGGTGTACAGGTGACCACCGCTGAGTACTGGGTGCGGCAGGTCCGGG 2700
V G D Q V T T A E Y W V R Q V R
ACACGCTCCGGTTCCGGCGAGCAGGTGGCCTCGTACGAGGACGCGGTGTTT 2750
D T V R F G E Q V A S Y E D A V F
30 GTCGAGCTGGGTGCCGACCGGTCACTGGCCCGCTGGTTCGACGGTGTGCG 2800
V E L G A D R S L A R L V D G V A
GATGCTGCACGGCGACCAAGAAATCCAGGCCGCGATCGGCGCCCTGGCCC 2850
M L H G D H E I Q A A I G A L A
ACCTGTATGTCAACGGCGTCACGGTTCGACTGGCCCGCTCCTGGGCGAT 2900
H L Y V N G V T V D W P A L L G D
35 CTTCCGGCAACACGGGTGCTGGACCTTCCGACATACGCCTTCCAGCACCA 2950
A P A T R V L D L P T Y A F Q H Q
GCGCTACTGGCTCGAGTCCGGTCCCCCGGCCACGGCCGACTCGGGCCACC 3000
R Y W L E S A P P A T A D S G H
CGTCTCGGCACCGGATCGCCGTGCGCGGTGCGCGGGCCGGGTGTTT 3050
40 F V L G T G V A V A G S P G R V F
ACGGGTCCCGTGGCCCGCGGTGCGGACCGCGCGGTGTTTCATCGCCGAAT 3100
T G P V P A G A D R A V F I A E L
GGCGCTCGCCCGCCGCGACGCCACCGACTGCGCCACGGTCAACAGCTCG 3150
A L A A A D A T D C A T V E Q L
45 ACGTCACCTCCGTGCCCCGCGGATCCGCCCGCGGACGGGCCACCGCGCAG 3200
D V T S V P G G S A R G R A T A Q
ACCTGGGTGCGATGAACCCCGCCGCGACGGGCGGCGCGCTTACCGTCCA 3250
T W V D E P A A D G R R R F T V H
CACCCGCGTCCGCGACGCCCCGTGGACGCTGCACGCCGAGGGGTTCTCC 3300
50 T R V G D A P W T L H A E G V L
GCCCCGGCCGCGTGGCCCGAGCCCGAAGCCGTGACACCGCCTGGCCCCCG 3350
R P G R V P Q P E A V D T A W P P
CCGGGCGCGGTGCCCCGCGGACGGGCTGCCCGGGCGTGGCGACGCGCGGA 3400
F G A V P A D G L P G A W R R A D
55 CCAGGTCTTCGTGGAAGCCGAAGTCGACAGCCCTGACGGCTTCGTGGCAC 3450
Q V F V E A E V D S P D G F V A
ACCCCGACCTGCTCGACGCGGTCTTCTCCGCGGTGCGGCGACGGGAGCCGC 3500
H P D L L D A V F S A V G D G S R
CAGCCGACCGGATGGCGCGACCTCGCGGTGCACGCGTCCGACGCCACCGT 3550
60 Q P T G W R D L A V H A S D A T V

GCTGCGCGCCTGCCTCACCCGCGCGGACAGTGGTGTGCTGGAGCTCGCCG 3600
 L R A C L T R R D S G V V E L A
 CCTTCGACGGTGCCGGAATGCCGGTGCTCACCCGCGGAGTCGGTGACGCTG 3650
 A F C G A G M P V L T A E S V T L
 5 GCGAGGTGCGGTGCGGACGCGGATCCGACGAGTCGGACGCTCTGCTTCG 3700
 G E V A S A G G S D E S D G L L R
 GCTTGAGTGGTTGCCGGTGCGGAGGCCACTACGACGGTGCCGACGAGC 3750
 L E W L P V A E A H Y D G A D E
 TGCCCGAGGGGTACACCCTCATCACCGCCACACACCCCGACGACCCCGAC 3800
 10 L P E G Y T L I T A T H F D C P D
 GACCCCAACCAACCCCAACACACCCACACGACCCCAACACAAACAC 3850
 D P T N P H N T P T R T H T Q T T
 ACGCGTCTTCACCGCCCTCCAACACCACCTCATCACCAACCAACACACCC 3900
 R V L T A L Q H H L I T T N H T
 15 TCATCGTCCACACCACCACCGACCCCGAGGCGCGCGCTCACCGGCCTC 3950
 L I V H T T T D P P G A A V T G L
 ACCCGCACCGGCACAAACGAACACCCCGGCGCATCCACCTCATCGAAAC 4000
 T R T A Q N E H P G R I H L I E T
 CCACACCCCAACCCCACTCCCTCACCCTCACCACCCCTCCACC 4050
 20 H H P H T P L P L T Q L T T L H
 AACCCCACTACGCTCACCAACAACACCTCCACACCCCACTCACC 4100
 Q P H L R L T N N T L H T P H L T
 CCCATCACCAACCACCAACACCACCAACACCCCAACACCCCAAC 4150
 P I T T H H N T T T T T P N T P P
 25 CCTCAACCCCAACCAACGCGCATCCTCATCACCGGCGGCTCCGGCACCCCTCG 4200
 L N P N H A I L I T G G S G T L
 CCGGCATCTCGCCCGCACCTCAACCAACCCCAACCTACCTCCTCTCC 4250
 A G I L A R H L N H P H T Y L L S
 CGCACACCAACCAACCCCAACCAACCCCGGACCCACATCCCTGCGACCT 4300
 30 R T P P P P T T P G T H I P C D L
 CACCGACCCCAACCAATCACCAAGCCCTCACCAACATACCAACCAACCC 4350
 T D P T Q I T Q A L T H I P Q P
 TCACCGGTCTCTTCACACCGCGCGCACCCCTCGACGAGCCACCTCACC 4400
 L T G I F H T A A T L D D A T L T
 35 AACCTCACCCCAACACCTCACCAACCAACCTCCAACCAACAGCGACGC 4450
 N L T P Q H L T T T L Q P K A D A
 CGCCTGGCACCTCCACCAACCAACCAACCAACCCCTCACCACTTCG 4500
 A W H L H H H T Q N Q P L T H F
 TCCTCTACTCCAGCGCGCGCCACCCCTCGGCAGCCCGGCAAGCCAAC 4550
 40 V L Y S S A A T L G S P G Q A N
 TACGCGCGCGCAACGCCTTCCTCGACGCGCTCGCCACCAACCGCCACAC 4600
 Y A A A N A F L D A L A T H R H T
 CCAAGGACAACCCGCCACCAACATCGCCTGGGGCATGTGGCACACCA 4650
 Q G Q P A T T I A W G M W H T T
 45 CCACACTCACCAACCACTCACCAACCAACCGGACCGGACCGCATCCGCGCG 4700
 T T L T S Q L T D S D R D R I R R
 GCGGCTTCCTGCGCATCTCGGACGACGAGGCGATGC
 G G F L P I S D D E G M

50 The *NheI*-*XhoI* hybrid FK-506 PKS module 8 containing the AT domain of module
 13 of rapamycin is shown below.

CCATGCGSCTGTACGAGGCGGCACGGCGCACCGGAAGTCCCGTGGTG 50
 M R L Y E A A R R T G S P V V V
 GCGGCCGCGCTCGACGACGCGCGGACGTGCCGCTGCTGCGCGGCTGCG 100
 A A A L D D A P D V P L L R G L R
 55 GCTACGACCGTCCGCGCTGCGCGCGTCCGGAACGCTCTCTCCCGGACC 150
 R T T V R R A A V R E R S L A D
 GCTCGCGGTGCTGCCGACGACGAGCGCGCGGACGCTCCCTCGCGTTG 200
 R S F C C P T T S A P T P P S R S

TCCTGGAACAGCACCGCCACCGTGCTCGGCCACCTGGGCGCCGAAGACAT 250
S W N S T A T V L G H L G A E D I
CCCGGCGACGACGTTCAAGGAACCTCGGCATCGACTCGCTCACCGCGG 300
F A T T T F K E L G I D S L T A
5 TCCAGCTGCGCAACGCGCTGACCACGGCGACCGGCGTACGCCTCAACGOC 350
V Q L R N A L T T A T G V R L N A
ACAGCGGTCTTCGACTTTCCGACGCGCGCGCGCTCGCCGCGAGACTCGG 400
T A V F D F P T P R A L A A R L G
CGACGAGCTGGCCCGTACCCGCGCGCCCGTTCGCGGCCCGGACCGCGGCCA 450
10 C E L A C T R A P V A A R T A A
CCCGGCGCCGCGACGACGAACCGCTGGCGATCGTGGGCATGGCCTGCCGT 500
T A A A H D E P L A I V G M A C R
CTCGCGGGCGGGGTCCCGTCGCCACAGGAGCTGTGGCGTCTCGTCGCGTC 550
L P G G V A S P Q E L W R L V A S
15 CGGCACCGACGCCATCACGGAGTTCGCCGCGGACCGCGGCTGGGACGTGG 600
G T D A I T E F P A D R G W D V
ACCGCTCTACGACCCCGACCCCGACGCGATCGGCAAGACCTTCGTCCGG 650
D A L Y D P D P D A I G K T F V R
CACGGCGGCTTCTCGACGGTGCGACCGGCTTCGACGCGGCGTTCTTCGG 700
20 H G G F L D G A T G F D A A F F G
GATCAGCCCGCGCGAGGCGCCTGGCCATGGACCCGACGCAACGGGTGCTCC 750
I S P R E A L A M D P Q Q R V L
TGGAGACGTCTCTGGGAGGCGTTTCGAAAGCGCGGGCATCACCCCGGACGCG 800
L E T S W E A F E S A G I T P D A
25 GCGCGGGGCAAGACACCGGCGTGTTCATCGGCGCGTTCTCTACGGGTA 850
A R G S D T G V F I G A F S Y G Y
CGGCACGGGTGCGGATACCAACGGCTTCGGCGCGACAGGGTCGCAGACCA 900
G T G A D T N G F G A T G S Q T
GCGTGCTCTCCGCGCGCCTCTCGTACTTCTACGGTCTGGAGGGCCCTTCG 950
30 S V L S G R L S Y F Y G L E G P S
GTCACGGTGCACACCGCCTGCTCGTCACTGGTCCGCTGCACCAGGC 1000
V T V D T A C S S S L V A L H Q A
AGGGCAGTCCCTGCGCTCGGGCGAATGCTCGCTCGCCCTGGTTCGGCGGTG 1050
G Q S L R S G E C S L A L V G G
35 TCACGGTGATGGCGTCCGCCGCGGATTTCGTTCGAGTTCTCCCGGCAGCGC 1100
V T V M A S P G G F V E F S R Q R
GGGCTCGCGCCGACGGGCGGGCGAAGGCGTTTCGGCGCGGGCGCGGACGG 1150
G L A P D G R A K A F G A G A D G
TACGAGCTTCGCGGAGGGCGCCGGTGCCCTGGTGGTTCGAGCGGCTCTCCG 1200
40 T S F A E G A G A L V V E R L S
ACGCGGAGCGCCACGGCCACACCGTCTCTGCCCTCGTACGCGGCTCCGCG 1250
D A E R H G H T V L A L V R G S A
GCTAACTCCGACGGCGCGTCAACGGTCTGTTCGGCGCCGAACGGCCCCCTC 1300
A N S D G A S N G L S A P N G P S
45 CCAGGAACGCGTCCATCCACGAGGCCCTCGGAACGCGAACTCACCCCGG 1350
Q E R V I H Q A L A N A K L T P
CCGATGTTCGACGCGGTTCGAGGCGCACGGCACCGGCACCCGCTCGGCGAC 1400
A D V D A V E A H G T G T R L G D
CCCATCGAGGCGCAGGCGCTGCTCGCGACGTACGGACAGGACCGGGCGAC 1450
50 P I E A Q A L L A T Y G Q D R A T
CCCCCTGCTGCTCGGCTCGCTGAAGTCAACATCGGGCACGCCAGGCCG 1500
F L L G S L K S N I G H A Q A
CGTCAGGGGTTCGCCGGATCATCAAGATGGTGCAGGCCATCCGGCACGGG 1550
A S G V A G I I K M V Q A I R H G
55 GAACTCCCGCGACACTGCACCGGACGAGCCGTTCGCCGACGTCCACTG 1600
E L P P T L H A D E P S P H V D W
TACCGCGGCTCGCGTCCAGCTCTGTACGTTCGCGCGGCGGTGGCGGGGA 1650
T A G A V E L T S A R P W F J
CCGGTTCGCCCGCGCGCTGCCGTCTCGTTCGTTTCGGCGTGAGCGGCACG 1700
60 T G R P R R A A V S S F G V S G T

SUBSTITUTE SHEET (RULE 26)

AACGCCCACATCATCCTTGAGGCAGGACCGGTCAAAACGGGACCGGTCTGA 1750
N A H I I L E A G P V K T G P V E
GGCAGGAGCGATCGAGGCAGGACCGGTCTGAAGTAGGACCGGTCTGAGGCTG 1800
A G A I E A G P V E V G P V E A
5 GACCGCTCCCCGCGGCGCCGCGCTCAGCACCGGGCGAAGACCTTCCGCTG 1850
G P L P A A P P S A P G E D L P L
CTCGTGTGCGGCGGTTCCTCCGAGGCACTCGACGAGCAGATCGGGCGCCT 1900
L V S A R S P E A L D E Q I G R L
10 GCGCGCCTATCTCGACACCGGCGCGGGCGTCTGACCGGGCGGCGCTGGCGC 1950
R A Y L D T G P G V D R A A V A
AGACACTGGCCCGGCGTACGCACTTCACCCACCGGGCCCTACTGCTCGGG 2000
Q T L A R R T H F T H R A V L L G
GACACCGTCTCGGCGCTCCCCCGCGGACCGAGGCGGACGAACCTCGTCTT 2050
D T V I G A P P A D Q A D E L V F
15 CGTCTACTCCGGTCTAGGGCACCCAGCATCCCGCGATGGGCGAGCAGCTAG 2100
V Y S G Q G T Q H P A M G E Q L
CCGATTCTGTCGGTGGTGTTCGCCGAGCGGATGGCCGAGTGTGCGGCGGCG 2150
A D S S V V F A E R M A E C A A A
TTGCGCGAGTTCGTGGACTGGGATCTGTTACGGTCTTGGATGATCCGGC 2200
20 L R E F V D W D L F T V L D D P A
GGTGGTGGACCGGTTGATGTGGTCCAGCCGCTTCTGGGCGATGATGG 2250
V V D R V D V V Q P A S W A M M
TTTCCCTGGCCGCGGTGTGGCAGGCGGCGGTGTGCGGCGGATGCGGTG 2300
V S L A A V W Q A A G V R P D A V
25 ATCGGCCATTTCGACGGGTGAGATCGCCGAGCTTGTGTGGCGGGTGCAGT 2350
I G H S Q G E I A A A C V A G A V
GTCACTACGCGATGCCGCGCGGATCGTGACCTTGGCGAGCCAGGCGATCG 2400
S L R D A A R I V T L R S Q A I
CCCCGGGCGCTGGCGGGCGGGGCGCGATGGCATCCGTGCGCCCTGCCCGCG 2450
30 A R G L A G R G A M A S V A L P A
CAGGATGTGAGCTGGTTCGACGGGGCGCTGGATCGCCGCCACAACGGGGC 2500
Q D V E L V D G A W I A A H N G P
CGCCTCCACCGTGATCGCGGGCACCCCGGAAGCGGTCTGACCATGTCTCA 2550
A S T V I A G T P E A V D H V L
35 CCGCTCATGAGGCACAAGGGGTGCGGGTGCGGCGGATCACCGTCTGACTAT 2600
T A H E A Q G V R V R R I T V D Y
GCCTCGCACACCCCGCACCTCGAGCTGATCCGCGACGAACCTACTCGACAT 2650
A S H T P H V E L I R D E L L D I
CACTAGCGACAGCTCGCAGACCCCGCTCGTGCCGTGGCTGTCCACCG 2700
40 T S D S S S Q T P L V P W L S T
TGGACGGCACCTGGGTCTGACAGCCCGCTGGACGGGGAGTACTGGTACCGG 2750
V D G T T W V D S P L D G E Y W Y R
AACCTGCGTGAACCGGTCTGGTTTCCACCCCGCGTCTAGCCAGTTGCAGGC 2800
N L R E P V G F H P A V S Q L Q A
45 CCAGGGCGACACCGTGTTCGTCTGAGGTCTAGCGCCAGCCCGGTGTGTTC 2850
Q G D T V F V E V S A S P V L L
AGGCGATGGACGACGATGTCGTACGGTTGCCACGCTGCGTCTGTGACGAC 2900
Q A M D D D V V T V A T L R R D D
GGCGACGCCACCGGATGCTCACCGCCCTGGCACAGGCCTATGTCCACGG 2950
50 G D A T R M L T A L A Q A Y V H G
CCTCACCGTCTGACTGGCCCGCCATCCTCGGCACCAACCAACCCGGGTAC 3000
V T V D W P A I L G T T T T R V
TGGACCTTCCGACCTACGCCCTTCCACACAGCGGTACTGGCTCGAGTCG 3050
L D L P T Y A F Q H Q R Y W L E S
55 GCTCCCCCGGCGGCGGACTCGGGCCACCCGTCCTCGGCACCGGAGT 3100
A P P A T A D S G H P V L G T G V
CGCCGTCGCGGGGTCTCGGGGCGGGGTGTTACGGGTCCCGTCCCCGCCG 3150
A V A G S P G R V F T G P V P A
GTGCGGACCGCGCGGTGTTTCATCGCCGAACCTGGCGCTCGCCGCCGCCGAC 3200
60 G A D R A V F I A E L A L A A A D

3250
A T D C A T V E Q L D V T S V P G
3300
S A R E R A T A Q T W V D E F
5
3350
A A D G R R R F T V H T R V G D A
3400
P W T L H A E G V L R P G R V P Q
10
3450
P E A V D T A W P P P G A V P A
3500
D G L P G A W R R A D Q V F V E A
3550
E V D S P D G F V A H P D L L D A
15
3600
V F S A V G D G S R Q P T G W R
3650
D L A V H A S D A T V L R A C L T
20
3700
R R D S G V V E L A A F D G A G M
3750
P V L T A E S V T L G E V A S A
3800
G G S D E S D G L L R L E W L P V
25
3850
A E A H Y D G A D E L P E G Y T L
3900
I T A T H P D D P D D P T N P H
3950
N T P T R T H T Q T T R V L T A L
30
4000
Q H A L I T T N H T L I V H T T T
4050
D P P G A A V T G L T R T A Q N
35
4100
E H P G R I H L I E T H H P H T P
4150
L P L Q L T T L H Q P H L R L T
4200
N N T L H T P H L T P I T T H H
40
4250
N T T T T T P N T P P L N P N H A
4300
I L I T G G S G T L A G I L A R H
45
4350
L N H P H T Y L L S R T P P P P
4400
T T P G T H I P C D L T D P T Q I
4450
T Q A L T H I P Q P L T G I F H T
50
4500
A A T L D D A T L T N L T P Q H
4550
L T T T L Q P K A D A A W H L E H
55
4600
H T Q N Q P L T H F V L Y S S A A
4650
A T L G S P G Q A N Y A A A N A
60
4700
F L D A L A T H R H T Q G Q P A T

ACCATCGCCTGGGGCATCTGGCACCACCACCACACTCACCAGCCAACT 4750
 T I A W G M W H T T T T L T S Q L
 CACCGACAGCGACCGCGACCGCATCCGCCGCGGGCTTCCTGCCGATCT 4800
 T C S C R C R I R R C G F L P I
 5 CCGACGACGAGGGCATGC
 S D D E G M

Example 3

Recombinant PKS Genes for 13-desmethoxy FK-506 and FK-520

10 The present invention provides a variety of recombinant PKS genes in addition to those described in Examples 1 and 2 for producing 13-desmethoxy FK-506 and FK-520 compounds. This Example provides the construction protocols for recombinant FK-520 and FK-506 (from *Streptomyces* sp. MA6858 (ATCC 55098), described in U.S. Patent Nos. 5,116,756, incorporated herein by reference) PKS genes in which the module 8 AT coding
 15 sequences have been replaced by either the *rapAT3* (the AT domain from module 3 of the rapamycin PKS), *rapAT12*, *eryAT1* (the AT domain from module 1 of the erythromycin (DEBS) PKS), or *eryAT2* coding sequences. Each of these constructs provides a PKS that produces the 13-desmethoxy-13-methyl derivative, except for the *rapAT12* replacement, which provides the 13-desmethoxy derivative, i.e., it has a hydrogen where the other
 20 derivatives have methyl.

Figure 7 shows the process used to generate the AT replacement constructs. First, a fragment of ~4.5 kb containing module 8 coding sequences from the FK-520 cluster of ATCC 14891 was cloned using the convenient restriction sites *SacI* and *SphI* (Step A in Figure 7). The choice of restriction sites used to clone a 4.0 - 4.5 kb fragment comprising
 25 module 8 coding sequences from other FK-520 or FK-506 clusters can be different depending on the DNA sequence, but the overall scheme is identical. The unique *SacI* and *SphI* restriction sites at the ends of the FK-520 module 8 fragment were then changed to unique *Bgl* II and *Nsi* I sites by ligation to synthetic linkers (described in the preceding Examples, see Step B of Figure 7). Fragments containing sequences 5' and 3' of the AT8
 30 sequences were then amplified using primers, described above, that introduced either an *Avr* II site or an *Nhe* I site at two different KS/AT boundaries and an *Xho* I site at the AT/DH boundary (Step C of Figure 7). Heterologous AT domains from the rapamycin and erythromycin gene clusters were amplified using primers, as described above, that introduced the same sites as just described (Step D of Figure 7). The fragments were ligated
 35 to give hybrid modules with in-frame fusions at the KS/AT and AT/DH boundaries (Step E of Figure 7). Finally, these hybrid modules were ligated into the *Bam* HI and *Pst* I sites of the

KC515 vector. The resulting recombinant phage were used to transform the FK-506 and FK-520 producer strains to yield the desired recombinant cells, as described in the preceding Examples.

The following table shows the location and sequences surrounding the engineered site of each of the heterologous AT domains employed. The FK-506 hybrid construct was used as a control for the FK-520 recombinant cells produced, and a similar FK-520 hybrid construct was used as a control for the FK-506 recombinant cells.

Heterologous AT	Enzyme	Location of Engineered Site
FK-506 AT8 (hydroxymalonyl)	<i>AvrII</i>	GGCCGT <u>ccgcgc</u> CGTGCGGCGGTCTCGTCGTTTC
	<i>NheI</i>	G R P R R A A V S S F ACCCAGCATCCCGCGATGGGTGAGCG <u>gctcgc</u> C
	<i>XhoI</i>	T Q H P A M G E R L A TACGCCTTCCAGCGGCGGCCCTACTGG <u>gtcgag</u> Y A F Q R R P Y W I E
rapamycin AT3 (methylmalonyl)	<i>AvrII</i>	GACCGG <u>ccccgt</u> CGGGCGGGCGTGTGTCCTTC
	<i>NheI</i>	D R P R R A G V S S F TGGCAGTGGCTGGGGATGGGCAGTGC <u>cctcgc</u> G
	<i>XhoI</i>	W Q W L G M G S A L R TACGCCTTCCAACACCAGCGGTACTGG <u>gtcgag</u> Y A F Q H Q R Y W V E
rapamycin AT12 (malonyl)	<i>AvrII</i>	GGCCGA <u>gcgcgc</u> CGGGCAGGCGTGTGTCCTTC
	<i>NheI</i>	G R A R R A G V S S F TCGCAGCGTGTGTCATGGGTGAGGA <u>actgac</u> C
	<i>XhoI</i>	S Q R A G M G E E L A TACGCCTTCCAGCACCAGCGCTACTGG <u>gtcgag</u> Y A F Q H Q R Y W L E
DEBS AT1 (methylmalonyl)	<i>AvrII</i>	GCGCGA <u>ccgcgc</u> CGGGCGGGGGTCTCGTCGTTTC
	<i>NheI</i>	A R P R R A G V S S F TGGCAGTGGGCGGGCATGGCCGTGCA <u>cctgct</u> C
	<i>XhoI</i>	W Q W A G M A V D L L TACCCGTTCCAGCGCGAGCGCGTCTGG <u>gtcgaa</u> Y P F Q R E R V W L E
DEBS AT2 (methylmalonyl)	<i>AvrII</i>	GACGGG <u>gtcgc</u> CGGGCAGGTGTGTCGGCGTTTC
	<i>NheI</i>	D G V R R A G V S A F GCCCAGTGGGAAGGCATGGCGCGGGA <u>gttgtt</u> G
	<i>XhoI</i>	A Q W E G M A R E L L TATCCTTTCCAGGGCAAGCGGTTCTGG <u>gtcgtg</u> Y P F Q G K R F W L L

The sequences shown below provide the location of the KS/AT boundaries chosen in the FK-520 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were engineered are indicated by lower case and underlining.

SUBSTITUTE SHEET (RULE 26)

CCGGCGCCGTCGAACTGCTGACGTCGGCCCCGGCGTGGCCCCGAGACCGACCGGccacggc
 A G A V E L L T S A R P W P E T D R P R
 GTGCGCGCGTCTCCTCGTTCGGGGTGAGCGGCACCAACGCCCACGTCATCCTGGAGGCGG
 P A A V S S F G V S G T N A H V T L E A
 5 GACCGGTAACGGAGACGCCCGCGGCATCGCCTTCCGGTGACCTTCCCTGCTGGTGTGCGG
 G P V T E T P A A S P S G D L P L L V S
 CACGCTCACCGGAAGCGCTCGACGAGCAGATCCGCCGACTGCGCGCCTACCTGGACACCA
 A R S P E A L D E Q I R R L R A Y L D T
 CCCCCGACGTCGACCGGGTGGCCGTGGCACAGACGCTGGCCCGGCGCACACACTTCGCCC
 10 T P D V D R V A V A Q T L A R R T H F A
 ACCGCGCCGTGCTGCTCGGTGACACCGTCATCACACACCCCCGCGGACCGGCCCCGACG
 H R A V L L G D T V I T T P P A D R P D
 AACTCGTCTTCGTCTACTCCGGCCAGGGCAGCCAGCATCCCGCGATGGGCGAGCagctcg
 E L V F V Y S G Q G T Q H P A M G E Q L
 15 CCGCGCCCATCCCGTGTTCGCCGACGCCTGGCATGAAGCGCTCCCGCCCTTGACAACC
 A A A H P V F A D A W H E A L R R L D N

The sequences shown below provide the location of the AT/DH boundary chosen in
 the FK-520 module 8 coding sequences. The region where an *XhoI* site was engineered is
 indicated by lower case and underlining.

TCCTCGGGGCTGGGTACGGCACGACCGCGATGTGCCCGCGTACGCGTTCCAACGGCGGC
 I L G A G S R H D A D V P A Y A F Q R R
 ACTACTGGatcgagTGGCACGCCCCGGCCGATCCGACGCGGGCCACCCCGTGTGGGCT
 25 H Y W I E S A R P A A S D A G H P V L G

The sequences shown below provide the location of the KS/AT boundaries chosen
 in the FK-506 module 8 coding sequences. Regions where *AvrII* and *NheI* sites were
 engineered are indicated by lower case and underlining.

TCGGCCAGGCGCGTGGCCGCGGACCGGCCGTccggcgcCGTGCGGCGGTCTCGTCTCGTTCGGG
 30 S A R P W P R T G R P R R A A V S S F G
 GTGAGCGGCACCAACGCCCACATCCTCTGGAGGCGGACCCGACCAGGAGGAGCCGTCG
 V S G T N A H I I L E A G P D Q E E P S
 GCAGAACCGGCCGGTGACCTCCCGCTGCTCGTGTGCGGACGGTCCCCGGAGGCACTGGAC
 A E P A G D L P L L V S A R S P E A L D
 35 GAGCAGATCGGGCGCCTGCGCGACTATCTCGACGCGCCCCCGCGTGGACCTGGCGGCC
 E Q I G R L R D Y L D A A P G V D L A A
 GTGGCGCGGACACTGGCCACGCGTACGCACTTCTCCACCGCGCCGTACTGCTCGGTGAC
 V A R T L A T R T H F S H R A V L L G D
 ACCGTCTACCCGCTCCCCCGTGGAAACAGCGGGCGAGCTCGTCTTCTGCTCTACTCGGGA
 40 T V I T A P P V E Q P G E L V F V Y S G
 CAGGGCACCCAGCATCCCGCGATGGGTGAGCGgctcgcCGCAGCCTTCCCCGTGTTCGCC
 Q G T Q H P A M G E R L A A A F P V F A
 GACCCGACGTACCCGCCTACGCCTTCCAGCGGCGGCCCTACTGGATCGAGTCCGCGCCG
 45 D P D V P A Y A F Q R R P Y W I E S A P

The sequences shown below provide the location of the AT/DH boundary chosen in
 the FK-506 module 8 coding sequences. The region where an *XhoI* site was engineered is
 indicated by lower case and underlining.

GACCCGACGTAACCGCCTACGCCTTCCAGCGGCGGCCCTACTGGatcgagTCCGCGCCG
 50 D P D V P A Y A F Q R R P Y W I E S A P

Example 4

Replacement of Methoxyl with Hydrogen or Methyl at C-15 of FK-506 and FK-520

The methods and reagents of the present invention also provide novel FK-506 and FK-520 derivatives in which the methoxy group at C-15 is replaced by a hydrogen or methyl. These derivatives are produced in recombinant host cells of the invention that
 5 express recombinant PKS enzymes the produce the derivatives. These recombinant PKS enzymes are prepared in accordance with the methodology of Examples 1 and 2, with the exception that AT domain of module 7, instead of module 8, is replaced. Moreover, the present invention provides recombinant PKS enzymes in which the AT domains of both modules 7 and 8 have been changed. The table below summarizes the various compounds
 10 provided by the present invention.

	Compound	C-13	C-15	Derivative Provided
	FK-506	hydrogen	hydrogen	13, 15-didesmethoxy-FK-506
	FK-506	hydrogen	methoxy	13-desmethoxy-FK-506
15	FK-506	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-506
	FK-506	methoxy	hydrogen	15-desmethoxy-FK-506
	FK-506	methoxy	methoxy	Original Compound -- FK-506
	FK-506	methoxy	methyl	15-desmethoxy-15-methyl-FK-506
	FK-506	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-506
20	FK-506	methyl	methoxy	13-desmethoxy-13-methyl-FK-506
	FK-506	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-506
	FK-520	hydrogen	hydrogen	13, 15-didesmethoxy FK-520
	FK-520	hydrogen	methoxy	13-desmethoxy FK-520
	FK-520	hydrogen	methyl	13,15-didesmethoxy-15-methyl-FK-520
25	FK-520	methoxy	hydrogen	15-desmethoxy-FK-520
	FK-520	methoxy	methoxy	Original Compound -- FK-520
	FK-520	methoxy	methyl	15-desmethoxy-15-methyl-FK-520
	FK-520	methyl	hydrogen	13,15-didesmethoxy-13-methyl-FK-520
	FK-520	methyl	methoxy	13-desmethoxy-13-methyl-FK-520
30	FK-520	methyl	methyl	13,15-didesmethoxy-13,15-dimethyl-FK-520

Example 5

Replacement of Methoxyl with Ethyl at C-13 and/or C-15 of FK-506 and FK-520

The present invention also provides novel FK-506 and FK-520 derivative compounds in which the methoxy groups at either or both the C-13 and C-15 positions are instead ethyl groups. These compounds are produced by novel PKS enzymes of the invention in which the AT domains of modules 8 and/or 7 are converted to ethylmalonyl specific AT domains by modification of the PKS gene that encodes the module.

Ethylmalonyl specific AT domain coding sequences can be obtained from, for example, the FK-520 PKS genes, the niddamycin PKS genes, and the tylosin PKS genes. The novel PKS genes of the invention include not only those in which either or both of the AT domains of modules 7 and 8 have been converted to ethylmalonyl specific AT domains but also those in which one of the modules is converted to an ethylmalonyl specific AT domain and the other is converted to a malonyl specific or a methylmalonyl specific AT domain.

Example 6

Neurotrophic Compounds

The compounds described in Examples 1 - 4, inclusive have immunosuppressant activity and can be employed as immunosuppressants in a manner and in formulations similar to those employed for FK-506. The compounds of the invention are generally effective for the prevention of organ rejection in patients receiving organ transplants and in particular can be used for immunosuppression following orthotopic liver transplantation.

These compounds also have pharmacokinetic properties and metabolism that are more advantageous for certain applications relative to those of FK-506 or FK-520. These compounds are also neurotrophic; however, for use as neurotrophins, it is desirable to modify the compounds to diminish or abolish their immunosuppressant activity. This can be readily accomplished by hydroxylating the compounds at the C-18 position using established chemical methodology or novel FK-520 PKS genes provided by the present invention.

Thus, in one aspect, the present invention provides a method for stimulating nerve growth that comprises administering a therapeutically effective dose of 18-hydroxy-FK-520. In another embodiment, the compound administered is a C-18,20-dihydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18-hydroxy-FK-520 derivative. In another embodiment, the compound administered is a C-13-desmethoxy and/or C-15-desmethoxy 18,20-dihydroxy-FK-520 derivative. In other embodiments, the compounds are the corresponding analogs of

FK-506. The 18-hydroxy compounds of the invention can be prepared chemically, as described in U.S. Patent No. 5,189,042, incorporated herein by reference, or by fermentation of a recombinant host cell provided by the present invention that expresses a recombinant PKS in which the module 5 DH domain has been deleted or rendered non-functional.

The chemical methodology is as follows. A compound of the invention (~200 mg) is dissolved in 3 mL of dry methylene chloride and added to 45 μ L of 2,6-lutidine, and the mixture stirred at room temperature. After 10 minutes, tert-butyldimethylsilyl trifluoromethanesulfonate (64 μ L) is added by syringe. After 15 minutes, the reaction mixture is diluted with ethyl acetate, washed with saturated bicarbonate, washed with brine, and the organic phase dried over magnesium sulfate. Removal of solvent *in vacuo* and flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) gives the protected compound, which is dissolved in 95% ethanol (2.2 mL) and to which is added 53 μ L of pyridine, followed by selenium dioxide (58 mg). The flask is fitted with a water condenser and heated to 70°C on a mantle. After 20 hours, the mixture is cooled to room temperature, filtered through diatomaceous earth, and the filtrate poured into a saturated sodium bicarbonate solution. This is extracted with ethyl acetate, and the organic phase is washed with brine and dried over magnesium sulfate. The solution is concentrated and purified by flash chromatography on silica gel (ethyl acetate: hexane (1:2) plus 1% methanol) to give the protected 18-hydroxy compound. This compound is dissolved in acetonitrile and treated with aqueous HF to remove the protecting groups. After dilution with ethyl acetate, the mixture is washed with saturated bicarbonate and brine, dried over magnesium sulfate, filtered, and evaporated to yield the 18-hydroxy compound. Thus, the present invention provides the C-18-hydroxyl derivatives of the compounds described in Examples 1 - 4.

Those of skill in the art will recognize that other suitable chemical procedures can be used to prepare the novel 18-hydroxy compounds of the invention. See, e.g., Kawai *et al.*, Jan. 1993, Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues, *FEBS Letters* 316(2): 107-113, incorporated herein by reference. These methods can be used to prepare both the C18-[S]-OH and C18-[R]-OH enantiomers, with the R enantiomer showing a somewhat lower IC₅₀, which may be preferred in some applications. See Kawai *et al.*, *supra*. Another preferred protocol is described in Umbreit and Sharpless, 1977, *JACS* 99(16): 1526-28, although it may be preferable to use 30 equivalents each of

SeO₂ and t-BuOOH rather than the 0.02 and 3-4 equivalents, respectively, described in that reference.

5 All scientific and patent publications referenced herein are hereby incorporated by reference. The invention having now been described by way of written description and example, those of skill in the art will recognize that the invention can be practiced in a variety of embodiments, that the foregoing description and example is for purposes of illustration and not limitation of the following claims.

Claims

1. An isolated nucleic acid that encodes a CoA ligase, a non-ribosomal peptide synthetase, or a domain of an extender module of a polyketide synthase enzyme that synthesizes FK-520.

5

2. The isolated nucleic acid of claim 1 that encodes an extender module, said module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

10

3. The isolated nucleic acid of claim 1 that encodes an open reading frame, said open reading frame comprising coding sequences for two or more extender modules, each extender module comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

15

4. The isolated nucleic acid of claim 1 that encodes a gene cluster, said gene cluster comprising two or more open reading frames, each of said open reading frames comprising coding sequences for two or more extender modules, each of said extender modules comprising a ketosynthase domain, an acyl transferase domain, and an acyl carrier protein domain.

20

5. The isolated nucleic acid of claim 2, wherein at least one of said domains is a domain of a module of a non-FK-520 polyketide synthase.

6. The isolated nucleic acid of claim 1, wherein said nucleic acid is a recombinant vector capable of replication in or integration into the chromosome of a host cell.

25

7. The isolated nucleic acid of claim 6 that is selected from the group consisting of cosmid pKOS034-120, cosmid pKOS034-124, cosmid pKOS065-M27, and cosmid pKOS065-M21.

30

8. The isolated nucleic acid of claim 5, wherein said non-FK-520 polyketide synthase is rapamycin polyketide synthase, FK-506 polyketide synthase, or erythromycin polyketide synthase.

9. A method of preparing a polyketide, said method comprising transforming a host cell with a recombinant DNA vector of claim 6, and culturing said host cell under conditions such that said polyketide synthase is produced and catalyzes synthesis of said polyketide.

10. The method of claim 9, wherein said host cell is a *Streptomyces* host cell.

11. The method of claim 9, wherein said polyketide is selected from the group consisting of FK-520, 13-desmethoxy-FK-520, and 13-desmethoxy-FK-506.

12. A recombinant host cell that expresses a recombinant polyketide synthase selected from the group consisting of: (i) an FK-520 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-520 polyketide synthase; (ii) an FK-506 polyketide synthase in which at least one AT domain is replaced by an AT domain of a non-FK-506 polyketide synthase; (iii) an FK-520 polyketide synthase in which at least one DH domain has been deleted; (iv) an FK-506 polyketide synthase in which at least one DH domain has been deleted.

13. The recombinant host cell of claim 12 that expresses an FK-520 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

14. The recombinant host cell of claim 12 that expresses an FK-506 polyketide synthase in which an AT domain of module 8 has been replaced by an AT domain that binds malonyl CoA, methylmalonyl CoA, or ethylmalonyl CoA.

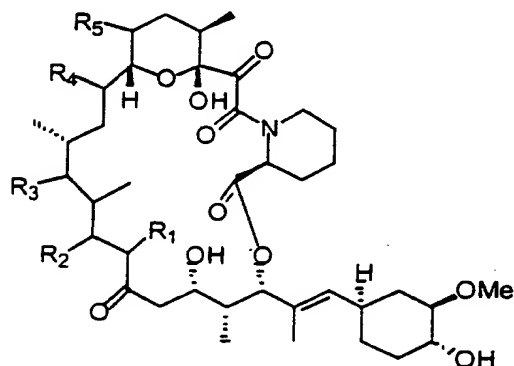
15. The recombinant host cell of claim 13, wherein a DH domain of module 5 or module 6 has been deleted.

16. The recombinant host cell of claim 14, wherein a DH domain of module 5 or module 6 has been deleted.

17. A recombinant host cell that comprises recombinant genes coding for enzymes sufficient for synthesis of ethylmalonyl CoA or 2-hydroxymalonyl CoA.

18. A polyketide having the structure

5



wherein, R₁ is hydrogen, methyl, ethyl, or allyl; R₂ is hydrogen or hydroxyl, provided that when R₂ is hydrogen, there is a double bond between C-20 and C-19; R₃ is hydrogen or hydroxyl; R₄ is methoxyl, hydrogen, methyl, or ethyl; and R₅ is methoxyl, hydrogen, methyl, or ethyl; but not including FK-506, FK-520, 18-hydroxy-FK-520, and 18-hydroxy-FK-506.

19. The polyketide of claim 18 that is 13-desmethoxy-FK-506.

15

20. The polyketide of claim 18 that is 13-desmethoxy-18-hydroxy-FK-520.

1 / 9

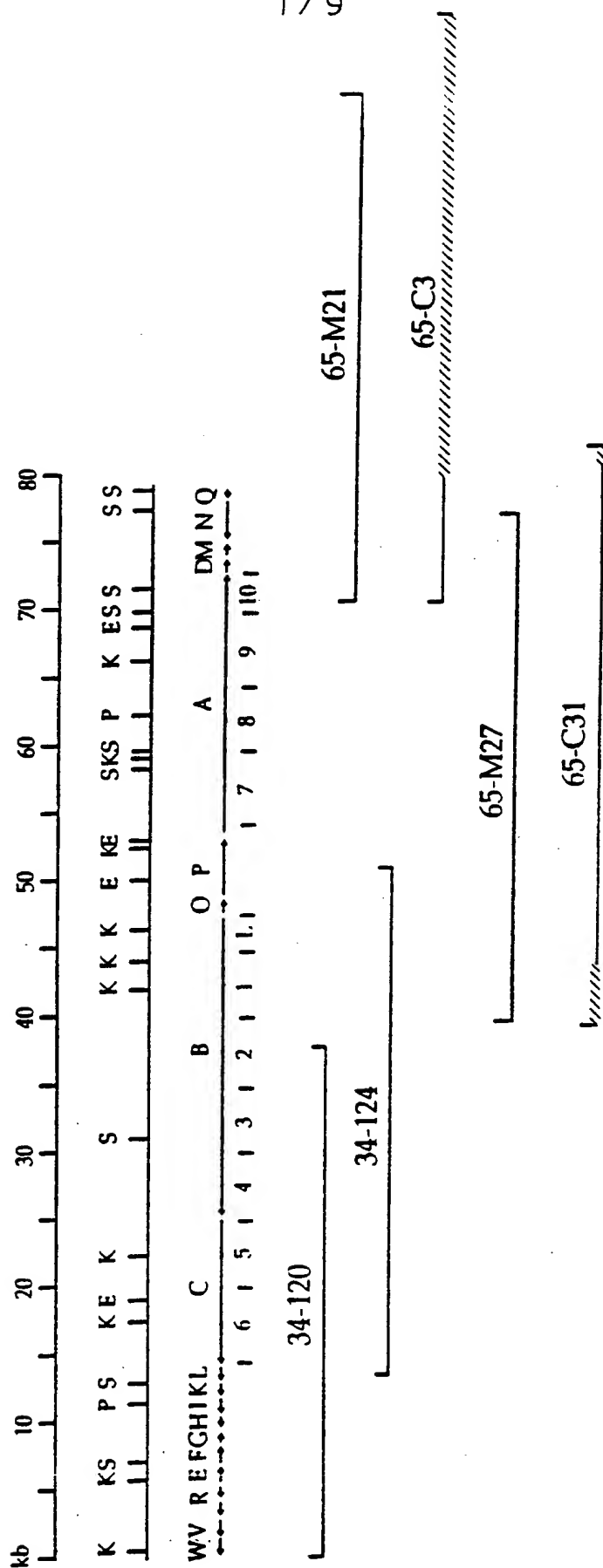


FIG. 1

2/9

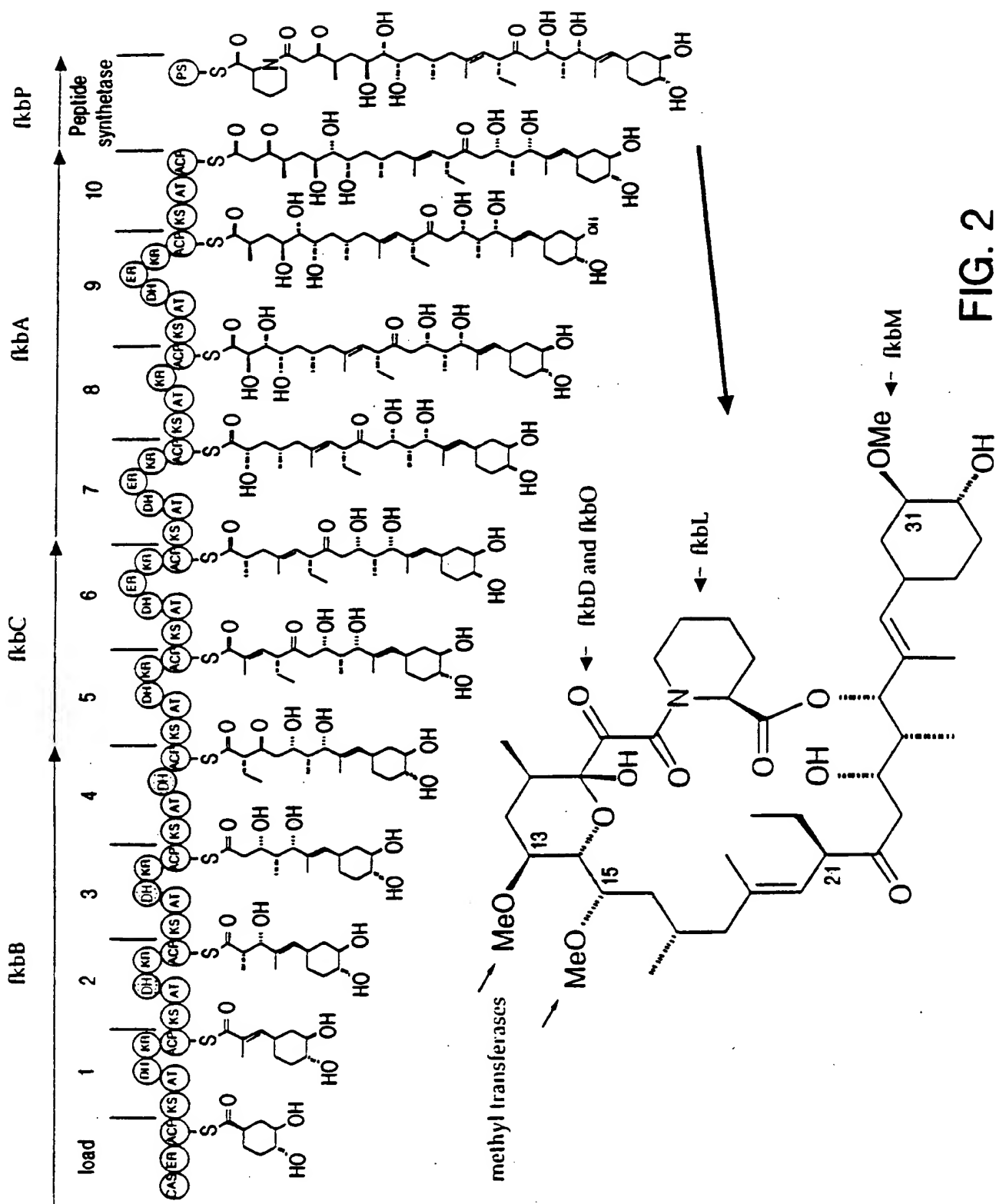


FIG. 2

SUBSTITUTE SHEET (RULE 26)

3/9

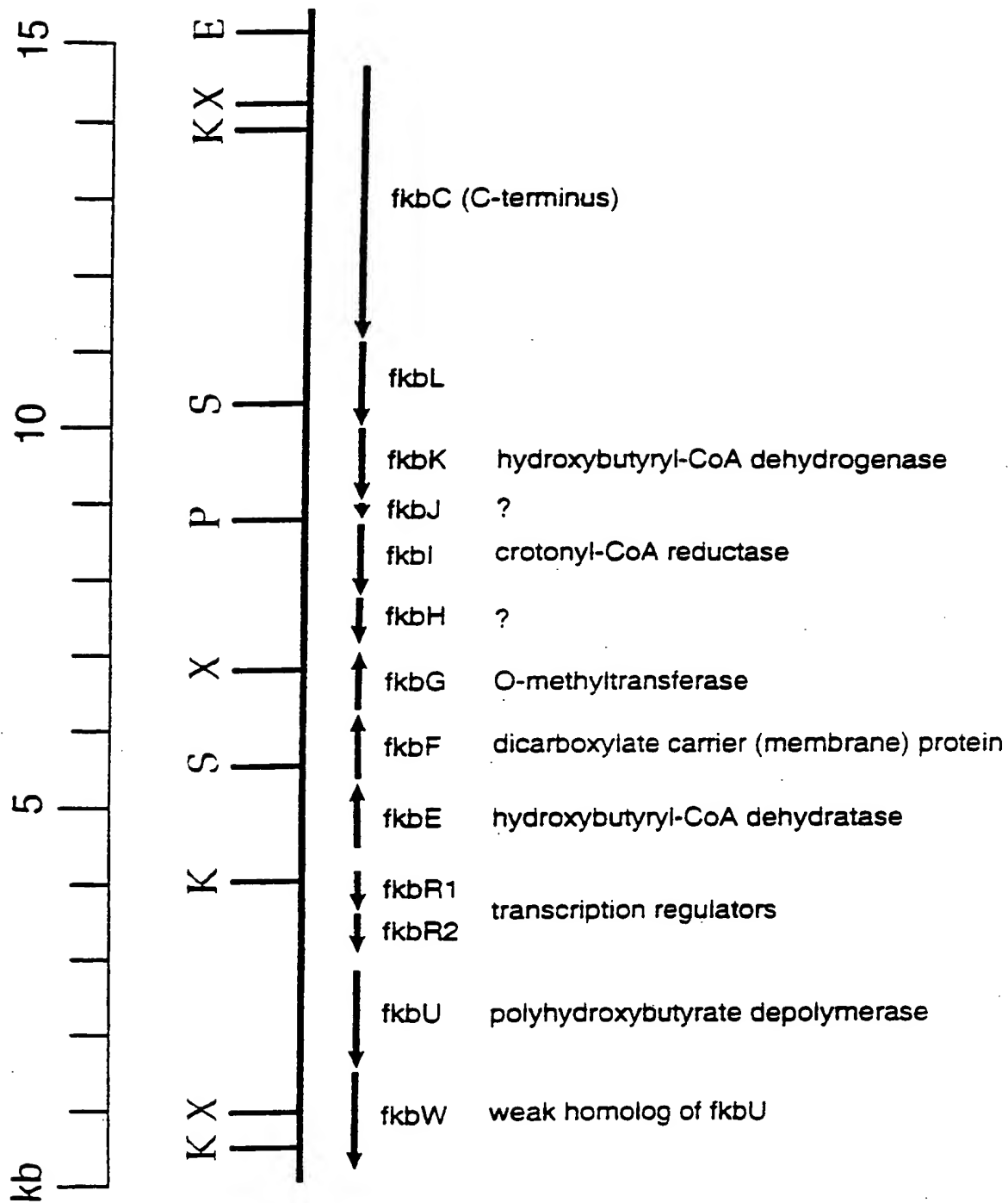


FIG. 3

SUBSTITUTE SHEET (RULE 26)

4 / 9

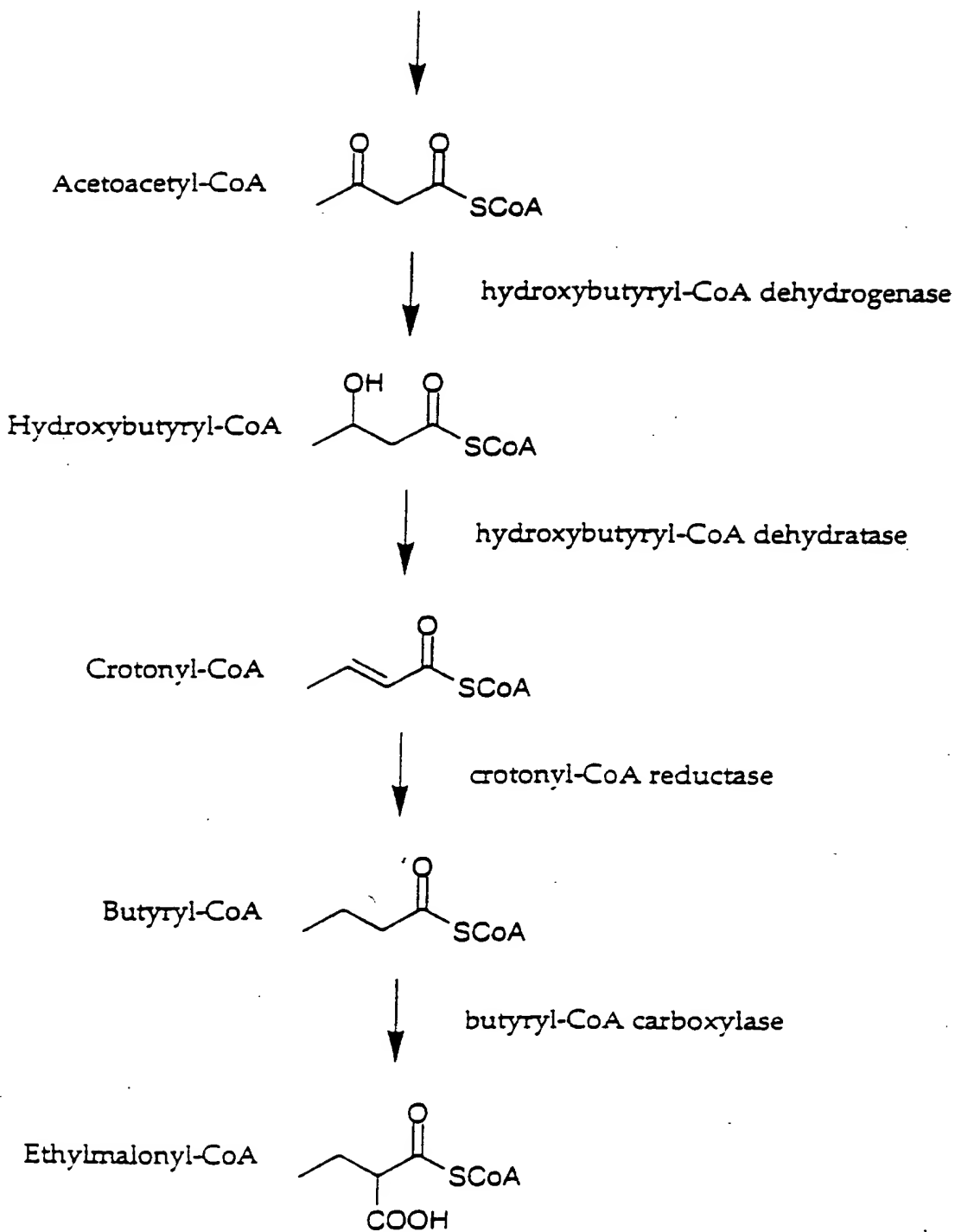
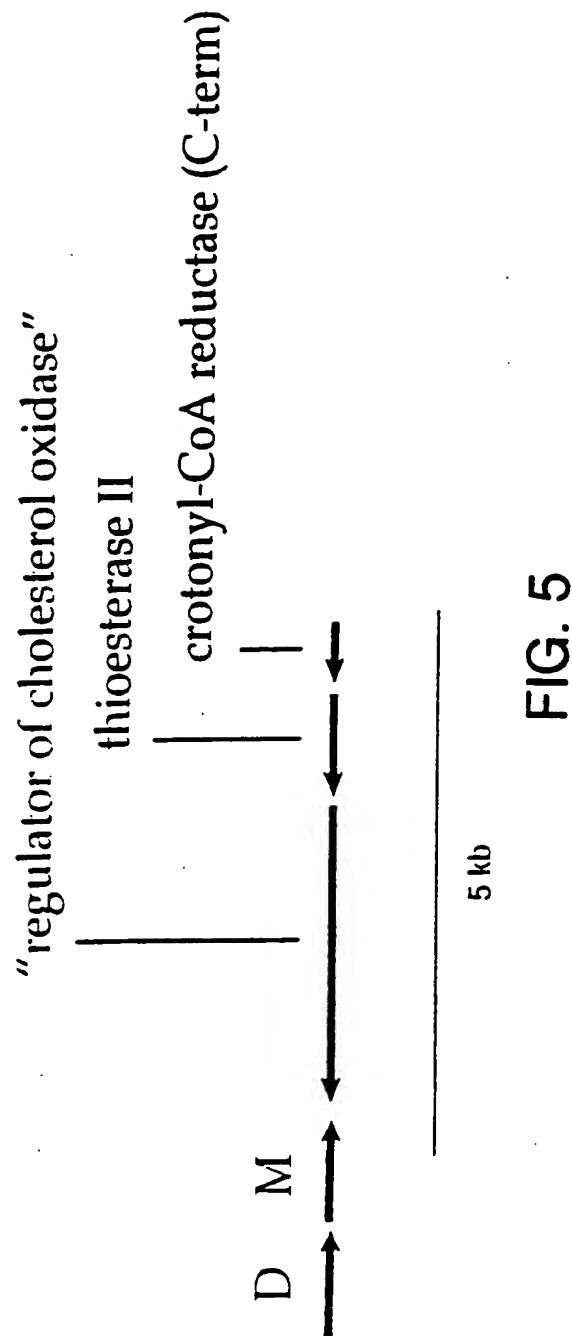


FIG. 4

SUBSTITUTE SHEET (RULE 26)



6/9

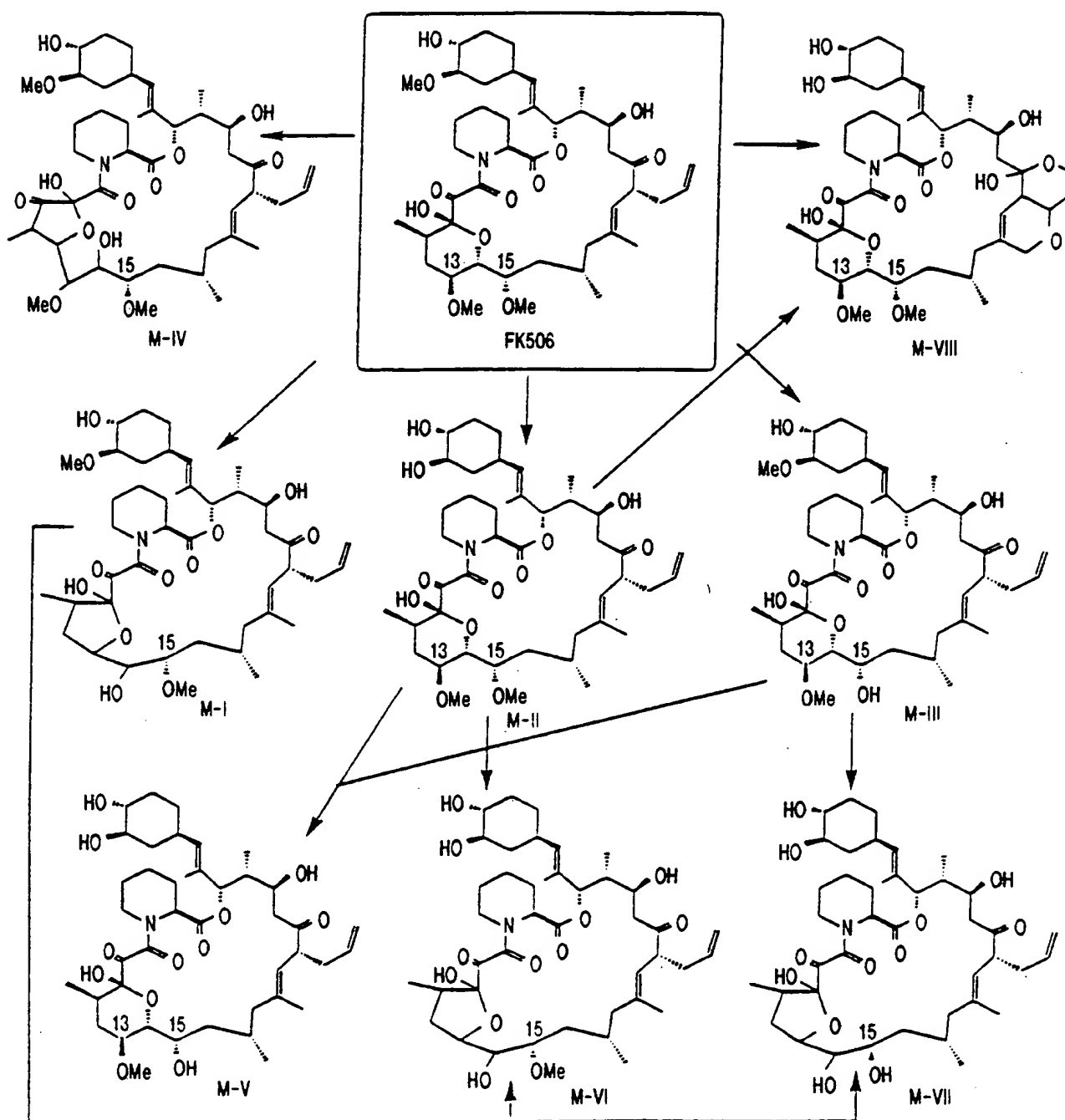


FIG. 6

SUBSTITUTE SHEET (RULE 26)

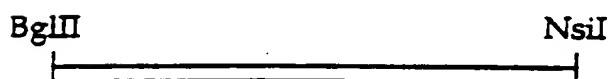
7 / 9

FIG. 7A



↓ linker insertion

FIG. 7B



↓ PCR amplification

FIG. 7C

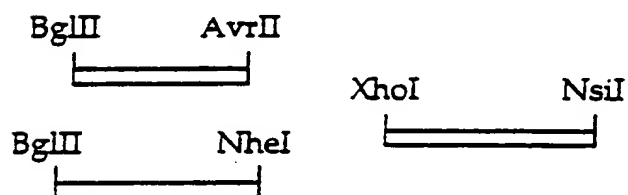
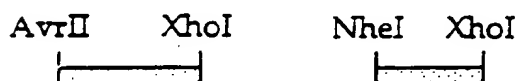
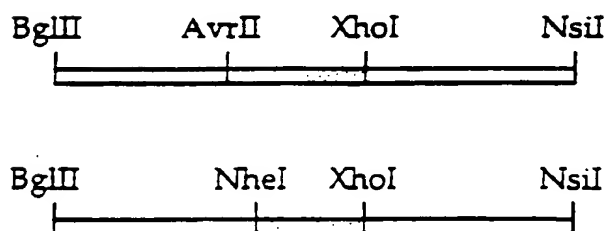


FIG. 7D



↓ ligation

FIG. 7E



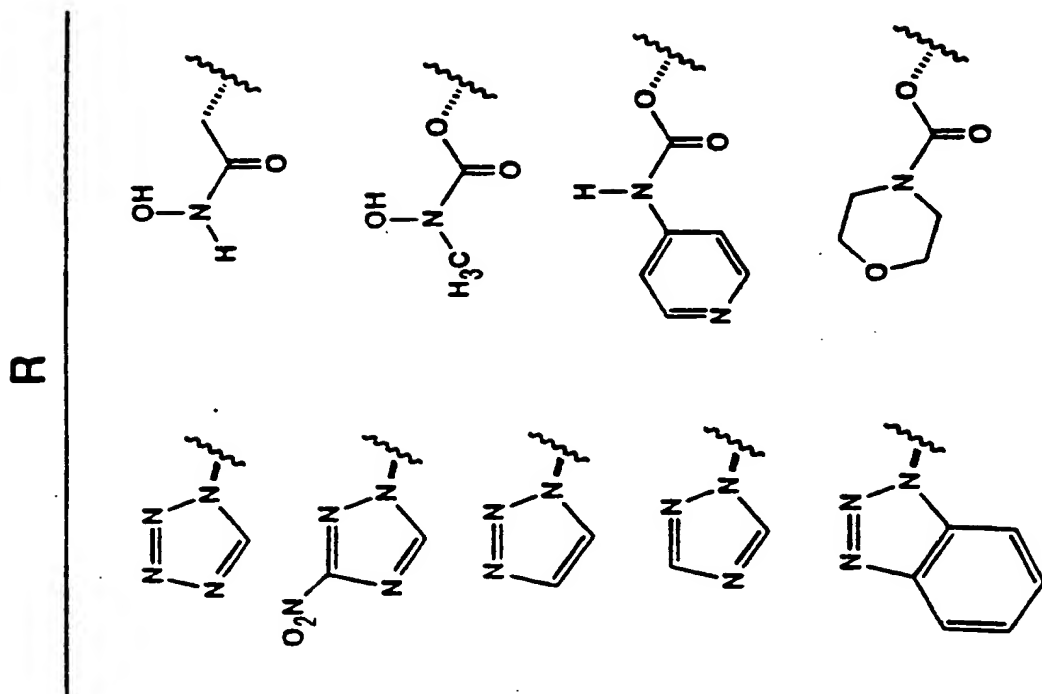
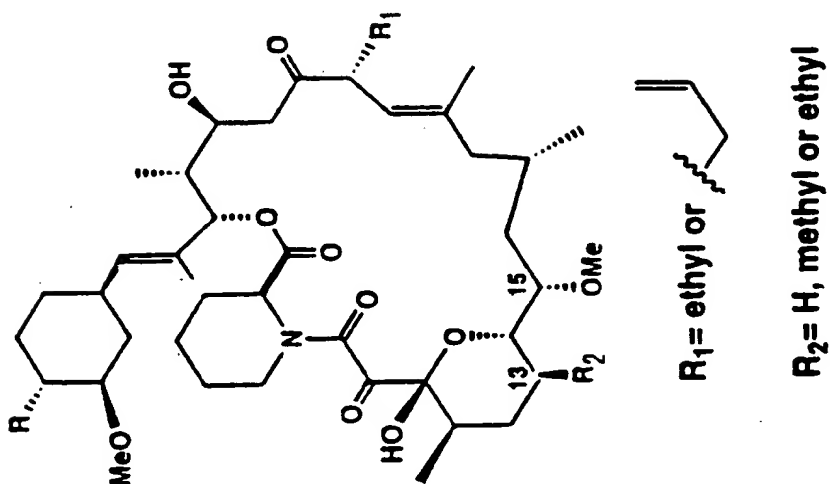


FIG. 8A



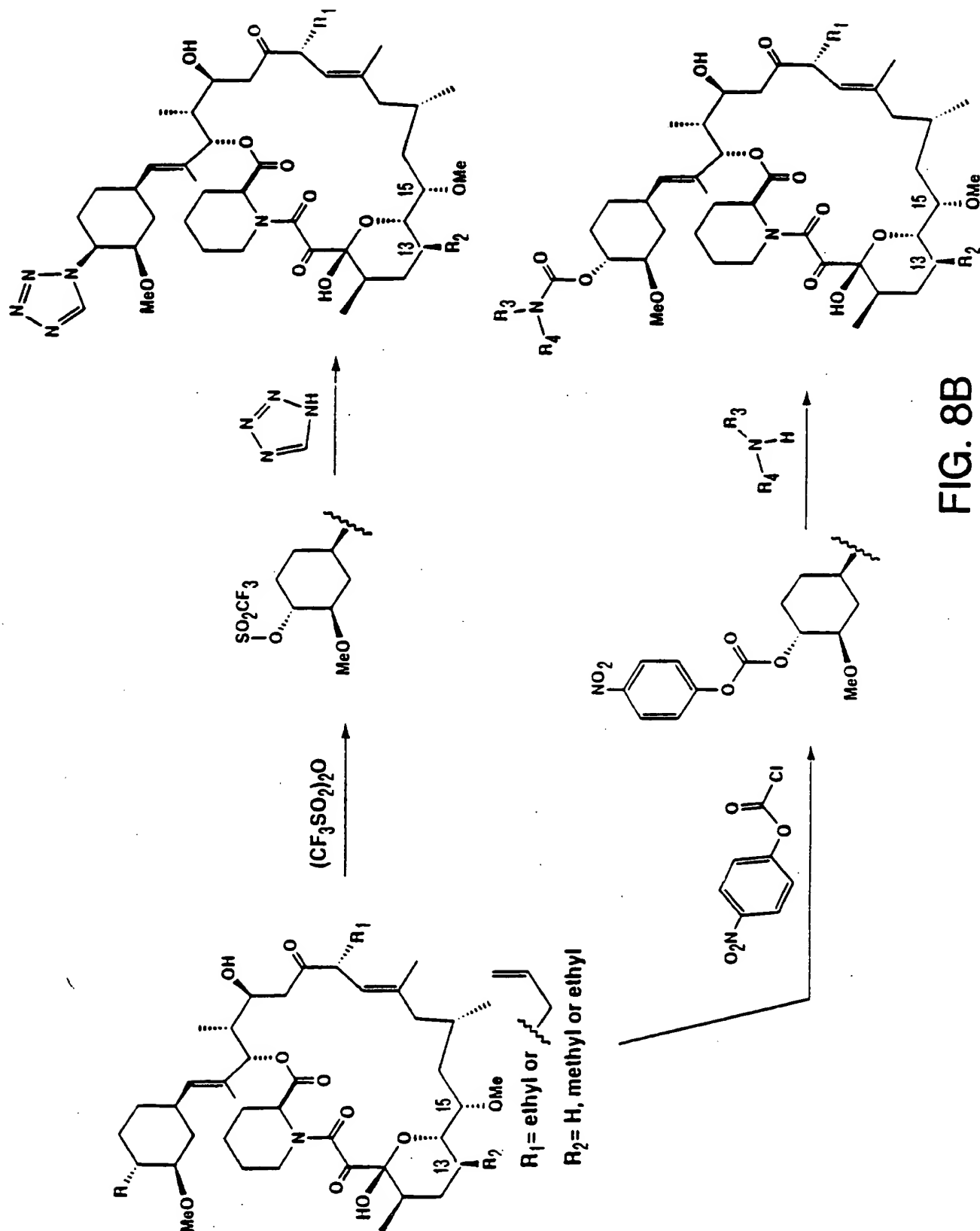


FIG. 8B

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

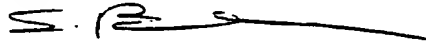
A. The indications made below relate to the microorganism referred to in the description on page <u>22</u> , line <u>31-33</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) <div style="text-align: center;">10801 University Blvd Manassas, VA 22110-2209 USA</div>	
Date of deposit 20 September 1999	Accession Number PTA-727, PTA-728 and PTA-729
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
All designated States	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	

<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input checked="" type="checkbox"/> This sheet was received by the International Bureau on: <div style="text-align: center; font-family: cursive; font-size: 1.2em;">13 JUN 00</div></div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer <div style="text-align: center; font-family: cursive; font-size: 1.2em;">S. B.</div></div>
---	---

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>22</u> , line <u>31-33</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <div style="text-align: center;">American Type Culture Collection</div>	
Address of depositary institution (including postal code and country) <div style="text-align: center;">10801 University Blvd Manassas, VA 22110-2209 USA</div>	
Date of deposit <div style="text-align: center;">20 September 1999</div>	Accession Number <div style="text-align: center;">PTA-726</div>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
<div style="text-align: center;">All designated States</div>	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
<div style="text-align: center;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">Authorized officer</div>	<div style="text-align: center;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px;"><input checked="" type="checkbox"/> This sheet was received by the International Bureau on: <div style="text-align: center; font-size: 1.2em;">13 JUN 00</div></div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">Authorized officer <div style="text-align: center;"></div></div>

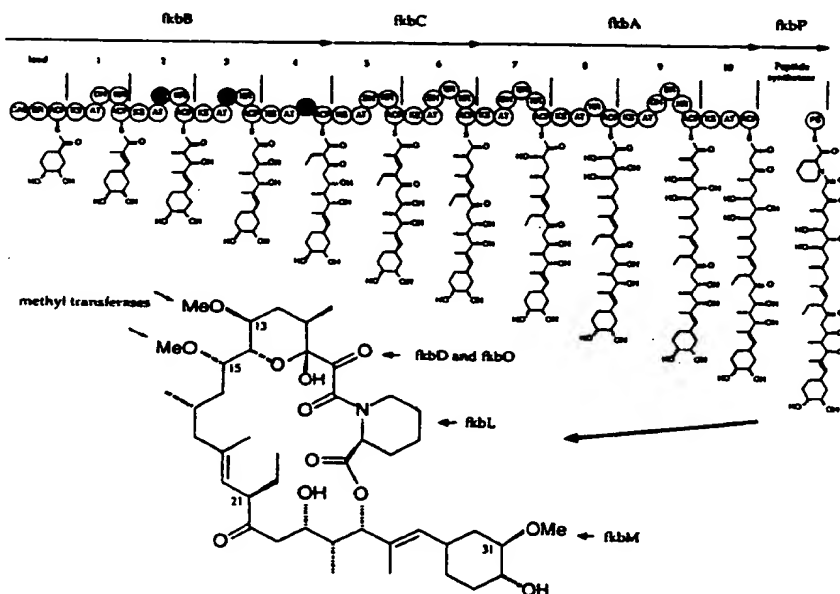
Form PCT/RO/134 (July 1992)



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/52, 15/54, 15/62, 9/10, C12P 17/18, 19/32, C07D 498/18 // (C07D 498/18, 311:00, 273:00, 211:00)		A3	(11) International Publication Number: WO 00/20601
(21) International Application Number: PCT/US99/22886		(43) International Publication Date: 13 April 2000 (13.04.00)	
(22) International Filing Date: 1 October 1999 (01.10.99)		(74) Agents: FAVORITO, Carolyn et al.; Morrison & Foerster LLP, 2000 Pennsylvania Avenue, N.W., Washington, DC 20006-1888 (US).	
(30) Priority Data: 60/102,748 2 October 1998 (02.10.98) US 60/123,810 11 March 1999 (11.03.99) US 60/139,650 17 June 1999 (17.06.99) US		(81) Designated States: AL, AM, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, DM, EE, GD, GE, HR, HU, IL, IS, JP, KG, KP, KR, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(71) Applicant (for all designated States except US): KOSAN BIOSCIENCES, INC. [US/US]; 3832 Bay Center Drive, Hayward, CA 94545 (US).		Published With international search report.	
(72) Inventors; and (75) Inventors/Applicants (for US only): REEVES, Christopher [US/US]; 4 East Altarinda Drive, Orinda, CA 94563 (US). CHU, Daniel [US/US]; 3767 Benton Street, Santa Clara, CA 95051 (US). KHOSLA, Chaitan [IN/US]; 740 Para Avenue, Palo Alto, CA 94306 (US). SANTI, Daniel [US/US]; 211 Belgrave Avenue, San Francisco, CA 94117 (US). WU, Kai [CN/US]; 900 Constitution Drive, Foster City, CA 94404 (US).			
(88) Date of publication of the international search report: 26 October 2000 (26.10.00)			

(54) Title: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA CONSTRUCTS THEREFOR



(57) Abstract

Host cells comprising recombinant vectors encoding the FK-520 polyketide synthase and FK-520 modification enzymes can be used to produce the FK-520 polyketide. Recombinant DNA constructs comprising one or more FK-520 polyketide synthase domains, modules, open reading frames, and variants thereof can be used to produce recombinant polyketide synthases and a variety of different polyketides with application as pharmaceutical and veterinary products.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/22886

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/52 C12N15/54 C12N15/62 C12N9/10 C12P17/18
C12P19/32 C07D498/18 //(C07D498/18,311:00,273:00,211:00)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12P C07D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, MEDLINE, STRAND, BIOSIS, EMBASE, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOTAMEDI H ET AL.: "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506" EUR. J. BIOCHEM., vol. 256, no. 3, 15 September 1998 (1998-09-15), pages 528-534, XP000906738 abstract figures 1,2 page 532, right-hand column, line 51 -page 533, left-hand column, line 18 --- -/--	12

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

27 July 2000

Date of mailing of the international search report

10/08/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

van de Kamp, M

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/22886

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	STASSI D L ET AL.: "Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering" PROC. NATL. ACAD. SCI. USA, vol. 95, no. 13, 23 June 1998 (1998-06-23), pages 7305-7309, XP002143632 abstract page 7308, left-hand column, line 4 -right-hand column, line 17 page 7309, left-hand column, line 24-40 ---	17
X	REYNOLDS K A ET AL.: "Rapamycin, FK506, and ascomycin -related compounds" DRUGS PHARM. SCI., vol. 82, 1997, pages 497-520, XP000906777 figure 3 page 502, line 7-25; figure 7 page 509-513, paragraph IV ---	18
X	DUMONT F J ET AL.: "The immunosuppressive and toxic effects of FK-506 are mechanistically related: pharmacology of a novel antagonist of FK-506 and rapamycin" JOURNAL OF EXPERIMENTAL MEDICINE, vol. 176, no. 3, 1 September 1992 (1992-09-01), pages 751-760, XP000906781 cited in the application abstract; figure 1 ---	18
X	KAWAI M ET AL.: "Structure-activity profiles of macrolactam immunosuppressant FK-506 analogues" FEBS LETTERS, vol. 316, no. 2, January 1993 (1993-01), pages 107-113, XP002143633 abstract scheme 1 table 1 ---	18
X	EP 0 323 042 A (FISONS PLC) 5 July 1989 (1989-07-05) example 13 ---	18
X	EP 0 356 399 A (SANDOZ AG ;SANDOZ AG (DE); SANDOZ AG (AT)) 28 February 1990 (1990-02-28) examples 2,3 ---	18
X	EP 0 463 690 A (MERCK & CO INC) 2 January 1992 (1992-01-02) example 3 ---	18
	-/--	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/22886

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
Y	MOTAMEDI H ET AL.: "Characterization of methyltransferase and hydroxylase genes involved in the biosynthesis of the immunosuppressants FK506 and FK520" J. BACTERIOLOGY, vol. 178, no. 17, July 1996 (1996-07), pages 5243-5248, XP002137077 abstract page 5245, left-hand column, line 1-3 figure 4	1-11
Y	--- MOTAMEDI H ET AL.: "Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK506" EUR. J. BIOCHEM., vol. 244, no. 1, 15 February 1997 (1997-02-15), pages 74-80, XP000906743 abstract page 79, left-hand column, line 26-35 page 75, left-hand column, line 31 -page 76, left-hand column, line 1	1-11
A	--- CHEN T S ET AL.: "Microbial transformation of immunosuppressive compounds. II. Specific desmethylation of 13-methoxy group of FK 506 and FR 9500520 by Actinomycete sp. ATCC 53828" J. ANTIBIOT., vol. 45, no. 4, April 1992 (1992-04), pages 577-580, XP002143634 figure 1	18-20
A	--- SHAFIEE A ET AL.: "Enzymatic synthesis and immunosuppressive activity of novel desmethylated immunomycins (ascomycins)" J. ANTIBIOT., vol. 46, no. 9, September 1993 (1993-09), pages 1397-1405, XP002143635 abstract	18,20
A	--- KHOSLA C: "Harnessing the biosynthetic potential of modular polyketide synthases" CHEMICAL REVIEWS, vol. 97, no. 7, 1997, pages 2577-2590, XP002130646 ISSN: 0009-2665 -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US-99/22886

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0323042 A	05-07-1989	AT 120466 T	15-04-1995
		AU 630866 B	12-11-1992
		AU 2822889 A	05-07-1989
		CA 1339128 A	29-07-1997
		CN 1033458 A	21-06-1989
		DE 3853477 D	04-05-1995
		DE 3853477 T	09-11-1995
		DK 387889 A	08-08-1989
		EP 0346427 A	20-12-1989
		ES 2071681 T	01-07-1995
		FI 90550 B	15-11-1993
		FI 930597 A	11-02-1993
		WO 8905304 A	15-06-1989
		IE 66163 B	13-12-1995
		IL 88629 A	12-04-1994
		JP 2502463 T	09-08-1990
		JP 2799208 B	17-09-1998
		NO 893166 A	04-08-1989
		NZ 227251 A	26-02-1990
		PT 89203 A, B	29-12-1989
		US 5376663 A	27-12-1994
		ZA 8809136 A	30-08-1989
EP 0356399 A	28-02-1990	AU 629563 B	08-10-1992
		AU 4024689 A	01-03-1990
		DK 418789 A	27-02-1990
		JP 2167287 A	27-06-1990
		NZ 230418 A	25-10-1991
		US 5011844 A	30-04-1991
		ZA 8906524 A	24-04-1991
EP 0463690 A	02-01-1992	CA 2044846 A	26-12-1991
		DE 69117221 D	28-03-1996
		DE 69117221 T	12-09-1996
		JP 4230288 A	19-08-1992
		US 5190950 A	02-03-1993
		US 5342935 A	30-08-1994

RECEIVED
MAY 13 2002
OPE/JCWS